

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 11:47:52 ; Search time 25 Seconds

(without alignments)
248,858 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798

Sequence: 1 MGVPNPNFSWVLPGRLAGLA.....GSIETFEDEKAVFQFQRTK 150

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	137.5	17.2	551	1	CC14_YEAST	Q00684 saccharomyc
2	119.5	15.0	989	1	PPP3_DICDI	P34637 dictyostell
3	115.5	14.5	663	1	DUS8_MOUSE	O09112 mus musculu
4	115	14.4	597	1	MCE1_HUMAN	O60942 homo sapien
5	115	14.4	597	1	MCE1_MOUSE	O55236 mus musculu
6	114.5	14.3	625	1	DUS8_HUMAN	O13202 homo sapien
7	114	14.3	168	1	PNP_NPYAC	P24656 autographa
8	114	14.3	665	1	DUSX_HUMAN	O9dy84 homo sapien
9	113	14.2	573	1	MCE1_CAEL	O17607 caenorhabd
10	111.5	14.0	340	1	DUSC_HUMAN	O9un16 homo sapien
11	107	13.4	220	1	PNP1_NPYOP	O10274 orgyia pseu
12	107	13.4	384	1	DUS9_HUMAN	O99956 homo sapien
13	104	13.0	276	1	PNP3_CHLEU	O33490 chlamydomon
14	104	13.0	595	1	PNP6_HUMAN	P23351 mus musculu
15	104	13.0	595	1	PNP6_MOUSE	O57668 methanococc
16	103.5	13.0	159	1	Y215_METUA	P23351 mus musculu
17	99	12.4	360	1	PNP7_HUMAN	P34236 homo sapien
18	98.5	12.3	699	1	PNPE_MOUSE	P44446 mus musculu
19	98	12.3	700	1	PNPE_HUMAN	P23469 homo sapien
20	96.5	12.1	147	1	Y220_METUA	O60280 methanococc
21	95.5	12.0	359	1	PNP7_RAT	P44446 mus musculu
22	94.5	11.8	280	1	DUS7_RAT	O63340 rattus norv
23	94.5	11.8	320	1	DUS7_HUMAN	O16829 homo sapien
24	94.5	11.8	320	1	DUS7_MOUSE	O91246 mus musculu
25	94.5	11.8	377	1	PNP2_DICDI	P34138 dictyostell
26	94	11.8	381	1	DUS6_HUMAN	O16828 homo sapien
27	94	11.8	381	1	DUS6_MOUSE	O96db1 mus musculu
28	94	11.8	381	1	DUS6_RAT	O64346 rattus norv
29	94	11.8	802	1	PNP8_MOUSE	P23352 mus musculu
30	93	11.7	458	1	PNP1_HUMAN	O99952 homo sapien
31	92.5	11.6	1705	1	PNP0_MOUSE	P70289 mus musculu
32	92	11.5	369	1	PNP5_RAT	P33234 rattus norv
33	92	11.5	432	1	PNP1_RAT	P20417 rattus norv

ALIGNMENTS

34	92	11.5	434	1	PTN1_CHICK	O13016 gallus gall
35	91.5	11.5	802	1	PTN1_HUMAN	P18433 homo sapien
36	91.5	11.5	829	1	PTN1_MOUSE	P18052 mus musculu
37	91.5	11.5	845	1	CSW_DROME	P29349 drosophila
38	91.5	11.5	1711	1	PTN2_RAT	O64612 rattus norv
39	91	11.4	913	1	PTN3_HUMAN	P26045 homo sapien
40	90.5	11.3	415	1	PTN2_HUMAN	P17706 homo sapien
41	90.5	11.3	541	1	PTN5_MOUSE	P54830 mus musculu
42	90	11.3	432	1	PTN1_MOUSE	P35821 mus musculu
43	90	11.3	489	1	MSG5_YEAST	P38590 saccharomyc
44	89.5	11.2	363	1	PTN2_RAT	P35233 rattus norv
45	89	11.2	403	1	PTN1_HUMAN	O00633 homo sapien

RESULT 1
ID CC14_YEAST
AC 000684;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable protein-tyrosine phosphatase CDC14 (EC 3.1.3.48).
GN CDC14 OR YFR028C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92283835; PubMed=1597462;
RA Wan J., Xu H., Grunstein M.;
RT "CDC14 of Saccharomyces cerevisiae. Cloning, sequence analysis, and
transcription during the cell cycle.";
RL J. Biol. Chem. 267:11274-11280(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasamura S.-I., Sasamura M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA MEDLINE=96287654; PubMed=8686381;
RA Eki T., Naitou M., Hagihara H., Ade M., Ozawa M., Sasamura S.-I.,
RA Sasamura M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,
RA Yamazaki M., Tashiro H., Hanaka F., Murakami Y.;
RT "Fifteen open reading frames in a 30.8 kb region of the right arm of
chromosome VI from Saccharomyces cerevisiae.";
RL Yeast 12:177-190(1996).
RN [4]
RP FUNCTION: PROBABLE PROTEIN-TYROSINE PHOSPHATASE THAT MAY BE
INVOLVED IN CHROMOSOME SEGREGATION.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
tyrosine + phosphate.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY IN
THE N-TERMINAL DUE TO A FRAMESHIFT.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

```

CC EMBL: M61194; AAA34477.1; ALT_FRAME.
DR EMBL: D50617; BAA09267.1; -.
DR PIR: A42784; A42784.
DR SGD: S0001924; CDC14.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSpC; 1.
DR SMART: SM00195; DSpC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Cell cycle; Hydrolyase.
FT ACT_SITE 283..283
SQ SEQUENCE 551 AA; 61906 MW; 4EB3985DFA3FDB23 CRC64;

Query Match 17.28; Score 137.5; DB 1; Length 551;
Best Local Similarity 28.0%; Pred. No. 2,9e-06;
Matches 30; Conservative 20; Mismatches 34; Indels 23; Gaps 3;

QY 27 HYGFLLDLGVRLVSLTERGPPHSDCPGLTLRLRLIPDCPPAPDQIDRFVQIVDEANNA 86
DB 238 HFE---DIGIQLDLIEDG-----TCDDLSI-----VKNFVGAETTIK 274

QY 87 RGEAVGVHCAIGFGRTGTMACYLVKERGLAAGDAIAEIRLRPGSI 133
DB 275 RGGKIAVHCAGKAGRTGCLIGAHILTYFTANECIGELFIRGMV 321

RESULT 2
PTP3_DICDI STANDARD; PRT; 989 AA.
ID PTP3_DICDI
AC P54637;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Protein-tyrosine phosphatase 3 (EC 3.1.3.48) (Protein-tyrosine-
DE phosphate phosphohydrolyase 3).
GN (PTP3 OR PTP3) AND (PTP2 OR PTP3).
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AX3;
RX MEDLINE=96189126; PubMed=8628311;
RA Gampert M., Howard P.K., Hunter T., Firtel R.A.;
RT "Multiple roles of the novel protein tyrosine phosphatase PTP3 during
RT Dictyostellium growth and development."
RL Mol. Cell. Biol. 16:2431-2444(1996).
CC -1- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 kDa (P130).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH
CC AND DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U38197; AAC47041.1; -.
DR HSP; 006124; ZSNP.
DR DictyDb; DD0111; PTPC1.
DR DictyDb; DD0777; PTPC2.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_PP.

```

```

DR Pfam: PF00102; Y-phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART: SM00194; PTPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolyase.
FT ACT_SITE 649..649
SQ SEQUENCE 989 AA; 109995 MW; 9371105AF80974AF CRC64;

Query Match 15.08; Score 119.5; DB 1; Length 989;
Best Local Similarity 28.0%; Pred. No. 0.00036;
Matches 37; Conservative 22; Mismatches 48; Indels 25; Gaps 5;

QY 33 DLGVR--HLVSLTERGPPHSDCPGLTLRLRLIPDF--CPAPDQIDRFVQIVDEANNA 89
DB 583 DIGVSLHLTKKGEEFREV-----VLLHTQWDCCAPSSIRRLSWVNTFKRGS 637

QY 90 A-----GVHCAIGFGRTGTMACYL-----VKEGLAAGDAIAEIRLRPGSI 132
DB 638 AKNTNGFVIVHCSAGISRGFTISINIMAKIERFGNDPSQMNISIDSVLELRORGM 697

QY 133 IETVEOKAVFO 144
DB 698 VQILDQYTFIFK 709

RESULT 3
DUS8_MOUSE STANDARD; PRT; 663 AA.
ID DUS8_MOUSE
AC O09112;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Neuronal tyrosine threonine phosphatase 1).
GN DUSP8 OR NTP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=96311565; PubMed=8733137;
RA Theodosiou A.M., Rodrigues N.R., Nesbitt M.A., Ambrose H.J.,
RA Paterson H., McEldown Arnold E., Boyd Y., Leversha M.A., Owen N.,
RA Blake D.J., Ashworth A., Davies K.E.;
RT "A member of the MAP kinase phosphatase gene family in mouse
RT containing a complex trinucleotide repeat in the coding region."
RL Hum. Mol. Genet. 5:675-684(1996).
CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN AND LUNG.

```

```

-1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC
CC -1- SIMILARITY: CONTAINS 1 RHODANESE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.sdb.ch).
CC
DR EMBL; X95518; CAA64772.1; .
DR HSSP; Q16828; IMRP.
DR MGPI; MG1:016626; NtPp1.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPc; 1.
DR SMART; SM00195; DSPc; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS0054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase; Nuclear protein.
KW DOMAIN 13
FT DOMAIN 162 432 RHODANESE.
FT DOMAIN 452 459 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 555 558 POLY-ARG.
FT DOMAIN 559 576 POLY-SER.
FT DOMAIN 577 600 POLY-GLY.
FT DOMAIN 311 552 POLY-SER.
FT ACT_SITE 246 246 PRO-RICH.
SQ SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;

Query Match 14.5%; Score 115.5; DB 1; Length 663;
Best Local Similarity 24.1%; Pred. NO. 0.00056;
Matches 41; Conservative 26; Mismatches 64; Indels 39; Gaps 6;

QY 8 FSWVLPGRLAG-----LALPLPLPA-----HYOFLDLGLVRLVSLTERGPP 48
Db 129 FSSCPGGLCEGPATLPMSLSQPLCLPVPVGLTRILPLHLVIGSOKDVLNMDIMQNGIS 188
QY 49 H-----SDSCPG-----LTLRLRLRP---DFCPAPQIDRFVOIYDEANARGANGVHCAL 97
Db 189 YVLANSNSNCPRDFICESRFMKIPINDMYCEKTLIPWLDKSIETFKALISSQVIVHCIA 248
QY 98 GFGFGTGLACLVYKERGLAAGDAIAETRLRLP-----GSTETVEQ 138
Db 249 GISRSATIAIAIYIMTKMGSSDDAIFRYKDRRPSISPFNPLGQLLEYER 298

RESULT 4
KCE1_HUMAN STANDARD; PRT; 597 AA.
AC 060942; Q43483; O60257; O60351;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE mRNA capping enzyme (HCE) (HCAPI) [Includes: Polynucleotide 5'-
DE triphosphatase (EC 3.1.3.33) (mRNA 5'-triphosphatase) (TPase); mRNA
DE guanylyltransferase (EC 2.7.7.50) (GMP--RNA guanylyltransferase)
DE (Gcase)].
GN RNGT OR CAP1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=96058741; PubMed=9371772;
RY Yue Z., Maldonado E., Pilliulla R., Cho H., Reinberg D., Shatkin A.J.;
RT "Mammalian capping enzyme complements mutant Saccharomyces cerevisiae

```

RT lacking RNA guanylyltransferase and selectively binds the elongating
RT form of RNA polymerase II." ;
RL Proc. Natl. Acad. Sci. U.S.A. 94:1298-12903(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), AND MUTAGENESIS.
RX MEDLINE=98181073; PubMed=9512541;
RA Yamada-Okabe T., Dol R., Shimmi O., Aritawa M., Yamada-Okabe H.;
RT "Isolation and characterization of a human cDNA for mRNA 5'-capping
RT enzyme." ;
RL Nucleic Acids Res. 26:1700-1706(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 4).
RC TISSUE=Colon adenocarcinoma;
RX MEDLINE=98139874; PubMed=9473487;
RA Tsukamoto T., Shibagaki Y., Murakoshi T., Suzuki M., Nakamura A.,
RT Gotoh H., Mizumoto K.;
RT "Cloning and characterization of two human cDNAs encoding the mRNA
RT capping enzyme." ;
RL Biochem. Biophys. Res. Commun. 243:101-106(1998).
CC -1- FUNCTION: BIFUNCTIONAL mRNA CAPPING ENZYME EXHIBITING RNA 5'-
CC GUANYLYLTRANSFERASE ACTIVITY IN THE N-TERMINAL PART AND MRNA
CC GUANYLYLTRANSFERASE ACTIVITY IN THE C-TERMINAL PART. CATALYZES THE
CC FIRST TWO STEPS OF CAP FORMATION: BY REMOVING THE GAMMA-PHOSPHATE
CC FROM THE 5'-TRIPHOSPHATE END OF NASCENT MRNA TO YIELD A
CC DIPHOSPHATE END, AND BY TRANSFERRING THE GMP MOIETY OF GTP TO THE
CC 5'-DIPHOSPHATE TERMINUS.
CC -1- CATALYTIC ACTIVITY: 5'-phosphopolynucleotide + H(2)O =
CC polynucleotide + phosphate.
CC -1- CATALYTIC ACTIVITY: GTP + (5')ppp-mRNA = diphosphate +
CC G(5')ppp-pur-mRNA.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; 1/HCEI/HCAP1A (SHOWN
CC HERE); 2/HCE1A; 3/HCE1B AND 4/HCAP1B; ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: ISOFORMS 1 AND 4 (AT A LESSER EXTENT) ARE
CC EXPRESSED IN CEREBRUM, CEREBELLUM, THYROID, LUNG, HEART, LIVER,
CC KIDNEY, SPLEEN, LARGE INTESTINE, TESTIS, SKIN AND MUSCLE.
CC -1- MISCELLANEOUS: ISOFORMS 2 TO 4 LACK MRNA 5'-GUANYLYLTRANSFERASE
CC ACTIVITY DUE TO DISRUPTIONS OF THE GSAFE DOMAIN.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION: BELONGS TO THE NON-RECEPTOR
CC CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION: BELONGS TO THE EUKARYOTIC
CC GSAFE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-ib.ch/announce/>
CC or send an email to licence@isb-ib.ch).
CC -----
DR EMBL; AF025654; AAB91559.1; -
DR EMBL; AB009022; BAA25894.1; -
DR EMBL; AB009023; BAA25895.1; -
DR EMBL; AB009024; BAA25896.1; -
DR EMBL; AB012142; BAA25198.1; -
DR EMBL; AB012143; BAA25199.1; -
DR Genew; HGNC:10073; RNCGT.
CC MIM; 603512; -
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR001339; mRNA_cap_enzyme.
DR Pfam; PF01331; mRNA_cap_enzyme; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
KW mRNA processing; mRNA capping; transferase; Nucleotidyltransferase;
KW Hydrolase; Multifunctional enzyme; Alternative splicing;
KW Nuclear protein.
FT DOMAIN 1 212 TPASE.
FT 229 597 GTASE.
FT DOMAIN 195 205 ASP/GLU-RICH.

FT	ACT SITE	126	126	RNA 5'-TRIPHOSPHATASE (BY SIMILARITY).
FT	ACT SITE	294	294	GUANYLATION SITE.
FT	VARSPIC	424	446	MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT	VARSPIC	481	597	MISSING (IN ISOFORM 3).
FT	VARSPIC	504	597	TREKQYDNKIKIECFENNNSWVFRQDRKSPNAYNTAMA
FT				VCSINSINPTVEEMTEFEEDICTAASOGQRKHHPDELM
FT				PEPPPRKRPPLT -> CLFIRSVLELDVLISIHONNANN
FT				OHISCSSTGG (IN ISOFORM 4).
FT	MUTAGEN	294	294	K->A: LOSS OF GTASE ACTIVITY.
FT	MUTAGEN	299	299	K->A: LOSS OF GTASE ACTIVITY.
FT	MUTAGEN	345	345	E->A: LOSS OF GTASE ACTIVITY.
FT	MUTAGEN	458	458	K->A: LOSS OF GTASE ACTIVITY.
FT	MUTAGEN	460	460	K->A: LOSS OF GTASE ACTIVITY.
FT	CONFLICT	30	30	M -> I (IN REF. 1).
FT	CONFLICT	484	484	Q -> P (IN REF. 1).
QY	SEQUENCE	597 AA;	68556 MW;	SICEECIBI90603DE CRC64;
QY	Query Match	14.4%;	Score 115;	DB 1; Length 597;
QY	Best Local Similarity	30.0%;	Pred. No. 0.00058;	
QY	Matches 24;	Conservative 18;	Mismatches 36;	Indels 2; Gaps 1
QY	69	PAPQIDREFOVIDEANARG--EAGVCHCALGFGFTGIMACYLKERGLAAGDAIAEIR	126	
QY	98	PTTNTTEFFILICERFENRPELIGVCHTGFNRTGFLICAFLEVKMDMSIAAVATPA	157	
QY	127	RLRGSITETYEQEKARPOFY	146	
QY	158	QARPPGIYKGDYLKELFERRY	177	
DB	158	QARPPGIYKGDYLKELFERRY	177	
RESULT 5				
MCEL_MOUSE				
ID	MCEL_MOUSE	STANDARD;	PRT;	597 AA.
AC	055236;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	mRNA capping enzyme (HCE) (MCEL) [Includes: polynucleotide 5'-			
DE	triphosphatase (EC 3.1.3.33) (mRNA 5'-triphosphatase) (trase); mRNA			
DE	(guanylyltransferase (EC 2.7.7.50) (GTP--RNA guanylyltransferase)			
DE	(trase)]			
GN	RNGT OR CAPLA.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteleota; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OC	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-294.			
RA	MEDLINE=98058741; Pubmed=9371772;			
RA	Yue Z., Maldonado E., Piluluta R., Cho H., Reinberg D., Shatkin A.J.;			
RA	"Mammalian capping enzyme complements mutant Saccharomyces cerevisiae			
RA	lacking RNA guanylyltransferase and selectively binds the elongating			
RT	form of RNA polymerase II.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 94:12898-12903(1997).			
RT	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=98070332; Pubmed=9407024;			
RA	McCracken S., Fong N., Rosonina E., Yankulov K., Brothers G.,			
RA	Siderovski D., Hessel A., Foster S., Shuman S., Bentley D.L.;			
RA	5'-capping enzymes are targeted to pre-mRNA by binding to the			
RT	phosphorylated carboxy-terminal domain of RNA polymerase II.";			
RT	Genes Dev. 11:3306-3318(1997).			
RT	[3]			
RP	CHARACTERIZATION, AND MUTAGENESIS.			
RA	MEDLINE=98445353; Pubmed=9770468;			
RA	Men Y., Yue Z., Shatkin A.J.;			
RT	"Mammalian capping enzyme binds RNA and uses protein tyrosine			
RT	phosphatase mechanism.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 95:12226-12231(1998).			
CC	-1- FUNCTION: BIFUNCTIONAL mRNA CAPPING ENZYME EXHIBITING RNA 5'-			
CC	TRIPHOSPHATASE ACTIVITY IN THE N-TERMINAL PART AND RNA			
CC	GUANYLYLTRANSFERASE ACTIVITY IN THE C-TERMINAL PART. CATALYZES THE			
CC	FIRST TWO STEPS OF CAP FORMATION: BY REMOVING THE GAMMA-PHOSPHATE			

	CC	FROM THE 5'-TRIPHOSPHATE END OF NASCENT RNA TO YIELD A DIPHOSPHATE END AND BY TRANSFERRING THE GMP MOIETY OF GTP TO THE 5-DIPHOSPHATE TERMINUS.
	CC	-1 CATALYTIC ACTIVITY: 5'-phosphopolynucleotide + H(2)O = polynucleotide + phosphate.
	CC	-1 CATALYTIC ACTIVITY: GTP + (5')pp-pur-mRNA = diphosphate + G(5')pp-pur-mRNA.
	CC	-1 ENZYME REGULATION: RNA TRIPHOSPHATASE ACTIVITY IS INHIBITED BY VANADATE, IODOACETATE, AND MAGNESIUM.
	CC	-1 SUBCELLULAR LOCATION: Nuclear.
	CC	-1 MISCELLANEOUS: BINDS TO THE ELONGATING PHOSPHORYLATED FORM OF RNA POLYMERASE II. CAN INTERACT INDIRECTLY BY BINDING TO POL. II C- TERMINAL DOMAIN AND DIRECTLY BY RNA BINDING. THE GRASP DOMAIN, RATHER THAN THE TPASE DOMAIN MEDIATES THESE INTERACTIONS. SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY. SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE EUKARYOTIC GRASP FAMILY.
	CC	-----
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its con- use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
	CC	-----
	DR	EMBL; AF025653; AA891558.1; .
	DR	EMBL; AF034568; AA888903.1; .
	DR	MGI; MGI:1329041; Raglt.
	DR	IrefPro; IPR000340;. DS phosphatase.
	DR	IrefPro; IPR000387; TYR_phosphatase.
	DR	IrefPro; IPR001339; mRNA_cap_enzyme.
	DR	Pfam; PF01331; mRNA_cap_enzyme. 1.
	DR	SMART; SMO0012; PTpc_DSPC. 1.
	DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
	DR	PROSITE; PS00566; TYR_PHOSPHATASE_2; 1.
	KM	mRNA processing; mRNA capping; transferase; Nucleotidyltransferase; Hydrolase; Multifunctional enzyme; Nuclear protein.
	FT	DOMAIN 1 212
	FT	DOMAIN 229 597
	FT	DOMAIN 195 205
	FT	ACT_SITE 126 126
	FT	ACT_SITE 294 294
	FT	ACT_SITE 36 36
	FT	MUTAGEN 66 66
	FT	MUTAGEN 110 110
	FT	MUTAGEN 125 125
	FT	MUTAGEN 126 126
	FT	MUTAGEN 132 132
	FT	MUTAGEN 133 133
	FT	MUTAGEN 138 138
	FT	MUTAGEN 168 168
	FT	MUTAGEN 290 290
	FT	MUTAGEN 294 294
	FT	MUTAGEN 315 315
	FT	MUTAGEN 315 315
	FT	MUTAGEN 530 530
	FT	MUTAGEN 530 530
	FT	MUTAGEN 533 533
	FT	MUTAGEN 533 533
	FT	MUTAGEN 537 537
	FT	MUTAGEN 537 537
	FT	MUTAGEN 537 537
	SEQ	SEQUENCE 597 AA; 68684 MW; BA9AE11889DEA7 CRC64;

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: L22858; AAA6631.1; -;
 DR EMBL: M6763; AAA6753.1; -;
 DR EMBL: M75679; -; NOT_ANNOTATED_CDS.
 DR PIR: C38477; C38477.
 DR PIR: A45431; A45431.
 DR PIR: A40781; A40781.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR SMART: SM00012; PRPC_DSPC; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 KW Hydrolyase; Late protein.
 FT ACT_SITE 119 119
 FT MOTAGEN 119 119
 FT CONFLICT 49 49
 FT CONFLICT 167 168
 SQ SEQUENCE 168 AA; 19288 MW; 2D4D85F1C52AB09 CRC64;

Query Match 14.3%; Score 114; DB 1; Length 168;
 Best Local Similarity 32.6%; Pred. No. 0.00017;
 Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;

QY 55 GLTHRLRIPDFCPADQIDRFQYIDFANAR--GENGVGCHALGSGRGTMCTACTYV 112
 DB 77 GLAKTKQVGQGTPEPSYQVEFDYKFTKCPGMLGVGHCHGNGINRGYMCRTLMH 136

QY 113 ERGLAAGDAIAEIRLRPGSIE--TYEOE 139
 DB 137 TLGIAPQEAIDRFKARGHKIERQNYOD 165

RESULT 8
 DUSX_HUMAN STANDARD; PRT; 665 AA.
 AC Q9BY84; Q9COG3;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dual specificity protein phosphatase 1? (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase phosphatase 7) (MKP-7).
 GN MKP7 OR KIAA1700.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21486429; PubMed=11489891;
 RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
 RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
 RT functions as a shuttle protein.";
 RL J. Biol. Chem. 276:39002-39011(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 CC -1- FUNCTION: Involved in the inactivation of MAP kinases.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +

CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 RHODANESE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AB052156; BAB40814.1; -;
 DR EMBL: AB051487; BAB21791.1; ALT_INTR.
 DR HSP: O16828; IMKP.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR001763; Rhodanese-1like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00581; Rhodanese; 2.
 DR Pfam: PF00782; DSPC; 2.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00012; PRPC_DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Nuclear protein.
 FT DOMAIN 11 131
 FT DOMAIN 227 289
 FT ACT_SITE 244 244
 FT ACT_SITE 244 244
 SQ SEQUENCE 665 AA; 73101 MW; 1BD853FE08460DF CRC64;

Query Match 14.3%; Score 114; DB 1; Length 665;
 Best Local Similarity 26.4%; Pred. No. 0.00083;
 Matches 28; Conservative 20; Mismatches 42; Indels 16; Gaps 3;

QY 50 SDSCPGTLT---HRLRIP---DFCPAPDQIDRFQYIDFANARGAVGVCHALGFGRT 102
 DB 192 SNCPKPDFIPESHFLAVPVNDSECEKILPWLDSVDFIEAKSNOCVILVHCLAGISRS 251

QY 103 GTMLACTYVKEGLAAGDAIAEIRLRP-----GSIEFYEOE 139
 DB 252 ATIAIAYIKRMKMSIDAEAYFVEKRPITSPNFNFGQLDYERK 297

RESULT 9
 MCEL_CAEEL STANDARD; PRT; 573 AA.
 ID MCEL_CAEEL
 AC Q17607; O02558;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE mRNA capping enzyme [includes: Polynucleotide 5'-triphosphatase
 DE (EC 3.1.3.33) (mRNA 5'-triphosphatase) (TPase); mRNA
 DE guanylyltransferase (EC 2.7.7.50) (GTP--RNA guanylyltransferase)
 DE (GTase)].
 GN CEL-1 OR C03D6.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Shuman S., Ho C.K.;
 RT "Identification of mRNA capping enzyme from C. elegans.";
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Burton J.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION, AND MUTAGENESIS OF CYS-124.
RX MEDLINE-97344078; Pubmed-9200605;
RA Takagi T., Moore C.R., Diehn F., Buratowski S.;
RT "An RNA 5'-triphosphatase related to the protein tyrosine
RT phosphatases."
RT Cell 89:867-873(1997).
CC -1- FUNCTION: BIFUNCTIONAL RNA CAPPING ENZYME EXHIBITING RNA 5'-
CC TRIPHOSPHATASE ACTIVITY IN THE N-TERMINAL PART AND RNA
CC GUANYLYLTRANSFERASE ACTIVITY IN THE C-TERMINAL PART. CATALYZES THE
CC FIRST TWO STEPS OF CAP FORMATION: BY REMOVING THE GAMMA-PHOSPHATE
CC FROM THE 5'-TRIPHOSPHATE END OF NASCENT RNA TO YIELD A
CC DIPHOSPHATE END, AND BY TRANSFERRING THE GMP MOIETY OF GTP TO THE
CC 5'-DIPHOSPHATE TERMINUS.
CC -1- CATALYTIC ACTIVITY: 5'-phosphopolynucleotide + H(2)O =
CC polynucleotide + phosphate.
CC -1- CATALYTIC ACTIVITY: GTP + (5')pp-pur-mRNA = diphosphate +
CC G(5')pp-pur-mRNA.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- INDUCTION: INHIBITED BY MAGNESIUM.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE NON-RECEPTOR
CC CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE EUKARYOTIC
CC GTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF003925; AB61344.1; -
DR EMBL: Z75525; CA89765.1; ALT_INIT.
DR Wormpep: C03D6.3; CE15578.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR001339; mRNA_cap_enzyme.
DR Pfam: PF01331; mRNA_cap_enzyme; 1.
DR SMART: SM00012; PTPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR mRNA processing; mRNA capping; Transference; Nucleotidyltransferase;
DR Hydrolyase; Multifunctional enzyme; Nuclear protein.
FT DOMAIN 1 212 TRASE.
FT FT 229 573
FT ACT_SITE 124 124 RNA 5'-TRIPHOSPHATASE.
FT ACT_SITE 299 299 GUANYLYLATION SITE (BY SIMILARITY).
FT MUTAGEN 124 124 C->S: LOSS OF ACTIVITY.
FT MUTAGEN 124 124 C->A: LOSS OF ACTIVITY.
SQ SEQUENCE 573 AA; 66336 MW; F50ACAS9FI815F47 CRC64;

Query Match 14.2%; Score 113; DB 1; Length 573;
Best Local Similarity 36.0%; Pred. No. 0.00088;
Matches 31; Conservative 12; Mismatches 37; Indels 6; Gaps 3;

DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Dual specificity protein phosphatase 12 (EC 3.1.3.48) (EC 3.1.3.16)
GN (Dual-specificity tyrosine phosphatase YVH1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE-99377030; Pubmed-10446167;
RA Nuda M., Manning E.R., Orth K., Dixon J.E.;
RT Identification of the human YVH1 protein-tyrosine phosphatase
RT orthologue reveals a novel zinc binding domain essential for in vivo
RT function".
RT J. Biol. Chem. 274:23991-23995(1999).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- COFACTOR: BINDS 2 MOLES OF ZINC PER MOLE OF PROTEIN.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS, HIGHEST EXPRESSION IN SPLEEN,
CC TESTIS, OVARY, AND PERIPHERAL BLOOD LEUKOCYTES AND LOWER
CC EXPRESSION IN LIVER AND LUNG.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF119226; AAD51134.1; -
DR HSSP: 016828; LMRP.
DR GeneW: HGNC:3067; DUSP12.
DR MIM: 604835; -
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00782; DSPC; 1.
DR SMART: SM00195; DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; FALSE NEG.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; UNKNOWN_1.
DR Hydrolyase; zinc; Metal-binding.
FT DOMAIN 88 160 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 115 115 BY SIMILARITY.
SQ SEQUENCE 340 AA; 37687 MW; 56B52192B42C73EB CRC64;

Query Match 14.0%; Score 111.5; DB 1; Length 340;
Best Local Similarity 28.4%; Pred. No. 0.00069;
Matches 40; Conservative 26; Mismatches 64; Indels 11; Gaps 6;

ID PTP1_NPVOB STANDARD; PRT; 220 AA.
 AC 010274;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Putative protein-tyrosine phosphatase 1 (EC 3.1.3.48).
 OS Orygia pseudotsugata multicaudata polyhedrosis virus (OPMVV).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 NC NCBI_TaxID=164623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9271300; PubMed=9126251;
 RA Atkins C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
 RA Rohmann G.F.;
 RT "The sequence of the Orygia pseudotsugata multicaudata nuclear
 RT polyhedrosis virus genome.";
 RL Virology 229:381-399(1997).
 CC -1- FUNCTION: COULD BE INACTIVE AS THE ACTIVE SITE CYSTEINE IS
 CC MODIFIED TO TRYPTOPHAN.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U75930; AAC5909.1; -
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR SMART; SM00012; PTPc_DSPc.1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR HYDROLASE.
 KW HYDROLASE.
 FT SITE 162 EQUIVALENT OF ACTIVE SITE CYS.
 FT SITE 220 AA; 25234 MW; D3FC093F195D45 CRC64;
 SQ SEQUENCE 220 AA; 25234 MW; D3FC093F195D45 CRC64;
 Query Match 13.4%; Score 107; DB 1; Length 220;
 Best Local Similarity 32.9%; Pred. No. 0.0012;
 Matches 27; Conservative 13; Mismatches 40; Indels 2; Gaps 1;
 QY 55 GLTHRLRIPFCPPAPDOIDRFVOYVDENARGEA--VGYHCALGRRGTMTLACTYLVK 112
 DB 120 GLTKKIRVPGRAVDDDIYAETVDEFFRRCPPTMLVAVWTHGLNRSGLVCRWAVE 179
 QY 113 ERGLAGDAIAEIRRLRPGSIE 134
 DB 180 RLGVSPDAIARFETANGHKIE 201
 RESULT 12
 DUS9_HUMAN STANDARD; PRT; 384 AA.
 AC 099956;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dual specificity protein phosphatase 9 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Mitogen-activated protein kinase phosphatase 4) (MAP kinase
 DE phosphatase 4) (MKP-4).
 GN DUSP9 OR MKP4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97184169; PubMed=9030581;
 RA Muda M., Boshett U., Smith A., Antonsom B., Gillieron C.,
 RA Chabert C., Camps M., Martinou I., Ashworth A., Arkinstall S.;
 RT "Molecular cloning and functional characterization of a novel
 RT mitogen-activated protein kinase phosphatase, MKP-4.";
 RL J. Biol. Chem. 272:5141-5151(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Platzer M.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INACTIVATES MAP KINASES. HAS A SPECIFICITY FOR THE ERK
 CC FAMILY.
 CC -1- CATALYTIC ACTIVITY: protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y08302; CA69610.1; -
 DR EMBL; U52111; AAF74507.1; -
 DR HSPB; Q16828; MKP.
 DR Genew; HGNC:3076; DUSP9.
 DR MIM; 300134; -
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPc.1.
 DR SMART; SM00195; DSPc.1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR HYDROLASE.
 KW DOMAIN 8 137 RHODANESE.
 FT DOMAIN 203 384 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 290 BY SIMILARITY.
 SQ SEQUENCE 384 AA; 41867 MW; F8598CA95AB379B7 CRC64;
 Query Match 13.4%; Score 107; DB 1; Length 384;
 Best Local Similarity 25.5%; Pred. No. 0.0022;
 Matches 36; Conservative 25; Mismatches 56; Indels 24; Gaps 4;
 QY 2 GVOPPNFS-----WVLPGRLAGIALPRLPAHYQFLDLGVRHVSITERGPHSDSCP 54
 DB 191 GATPPVGLRASFPVGLPNNLYLSA--RDSANESLSAKIGIRIILWATLPNPFEEKNG 248
 QY 55 GLTHRLRIPD-----PCPPAPDOIDRFVOYVDENARGEAVGACGFGRTGTM 106
 DB 249 DFHYKOIPDISDHSQNSLRFEPEA-----IEFIDBALSONCGVLVHCLAGVRSYTVT 301
 QY 107 ACYLVKERGLAGDAIAEIR 127
 DB 302 VAVIMQKTLHLNDAYDLVKR 322
 RESULT 13
 PTP3_CHLEU STANDARD; PRT; 276 AA.
 ID PTP3_CHLEU
 AC 039491;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Dual specificity protein phosphatase (EC 3.1.3.48) (EC 3.1.3.16).
 GN VH-PTP13.
 OS Chlamydomonas eugametos.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadales; Chlamydomonadales;
 OC Chlamydomonadales; Chlamydomonadales;
 OC NCBI_Taxid=3053;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-UTEX 10;
 RC MEDLINE=95323001; PubMed=7599654;
 RA Haring M.A., Siderius M., Jonak C., Hirt H., Walton K.M.,
 RA Musgrave A.;
 RT "Tyrosine phosphatase signalling in a lower plant: cell-cycle and
 RT oxidative stress-regulated expression of the Chlamydomonas eugametos
 RT VH-PTP13 gene.";
 RL Plant J. 7:981-988(1995).
 CC -1- FUNCTION: COULD BE INVOLVED IN TYROSINE PHOSPHATASE SIGNALLING
 CC PATHWAYS, HAVING MAP-KINASES AS SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- DEVELOPMENTAL STAGE: NONDIVIDING GAMETES DID NOT EXPRESS THE VH-
 CC PTP13 GENE WHEREAS SYNCHRONOUSLY DIVIDING VEGETATIVE CELLS ONLY
 CC EXPRESSED VH-PTP13 IN THE EARLY G1-PHASE OF THE CYCLE.
 CC -1- INDUCTION: BY OXIDATIVE STRESS.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; X77938; CA54910.1; -
 CC DR HSSB; P51452; 1VHR.
 CC DR InterPro; IPR000340; DS_Phasphatase.
 CC DR InterPro; IPR000387; TYR_Phasphatase.
 CC DR Pfam; PF00782; DSPC; 1.
 CC DR SMART; SM00195; DSPC; 1.
 CC DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 CC DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 CC DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
 CC KM Hydrolase.
 CC FT DOMAIN 89 276 CATALYTIC.
 CC FT ACT_SITE 172 172
 CC FT MUTAGEN 172 172 C->S: INACTIVE.
 CC SQ SEQUENCE 276 AA; 30310 MW; 034EF6951E03381 CRC64;
 CC
 CC Query Match 13.0%; Score 104; DB 1; Length 276;
 CC Best Local Similarity 27.6%; Pred. No. 0.0031;
 CC Matches 35; Conservative 21; Mismatches 57; Indels 14; Gaps 4;

RESULT 14
 PTH6_HUMAN

ID PTH6_HUMAN STANDARD; PRT; 595 AA.
 AC P29350;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Protein-tyrosine phosphatase, non-receptor type 6 (EC 3.1.3.48)
 DE (Protein-tyrosine phosphatase 1C) (PTP-1C) (Hematopoietic cell
 DE protein-tyrosine phosphatase) (SH-PTP1) (Protein-tyrosine phosphatase
 DE SHP-1).
 GN PTPN6 OR PTP1C OR HCP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE=92123209; PubMed=1732748;
 RA Y.T., Cleveland J.L., Ihle J.N.;
 RT "Protein tyrosine phosphatase containing SH2 domains:
 RT characterization, preferential expression in hematopoietic cells, and
 RT localization to human chromosome 12p12-p13.";
 RL Mol. Cell. Biol. 12:836-846(1992).
 RN (2)
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC TISSUE-Breast;
 RX MEDLINE=91343005; PubMed=1652101;
 RA Shen S.H., Baestien L., Posner B.I., Chretien P.;
 RT "A protein-tyrosine phosphatase with sequence similarity to the SH2
 RT domain of the protein-tyrosine kinases.";
 RL Nature 352:736-739(1991).
 RN (3)
 RP REVISIONS.
 RA Shen S.H., Baestien L., Posner B.I., Chretien P.;
 RL Nature 353:868-868(1991).
 RN (4)
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE=92141214; PubMed=1736296;
 RA Plutsky J., Neel B.G., Rosenberg R.D.;
 RT "Isolation of a src homology 2-containing tyrosine phosphatase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:1123-1127(1992).
 RN (5)
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=95394454; PubMed=7665165;
 RA Banville D., Stocco R., Shen S.H.;
 RT "Human protein tyrosine phosphatase 1C (PTPN6) gene structure:
 RT alternate promoter usage and exon skipping generate multiple
 RT transcripts.";
 RL Genomics 27:165-173(1995).
 RN (6)
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=96303695; PubMed=8723724;
 RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
 RA Malley T., Gibbs R.A.;
 RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
 RT genes at human chromosome 12p13.";
 RL Genome Res. 6:314-326(1996).
 RN (7)
 RP PHOSPHORYLATION.
 RX MEDLINE=95300784; PubMed=7781604;
 RA Li R.Y., Galt F., Ragab A., Ragab-Thomas J.M.F., Chap H.;
 RT "Tyrosine phosphorylation of an SH2-containing protein tyrosine
 RT phosphatase is coupled to platelet thrombin receptor via a pertussis
 RT toxin-sensitive heterotrimeric G-protein.";
 RL EMBO J. 14:2519-2526(1995).
 RN (8)
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 248-399.
 RX MEDLINE=98447672; PubMed=9774441;
 RA Yang J., Liang X., Niu T., Meng W., Zhao Z., Zhou G.W.;
 RT "Crystal structure of the catalytic domain of protein-tyrosine
 RT phosphatase SHP-1.";
 RL J. Biol. Chem. 273:28199-28207(1998).
 CC -1- FUNCTION: PLAYS A KEY ROLE IN HEMATOPOIESIS. THIS PTPASE ACTIVITY
 CC MAY DIRECTLY LINK GROWTH FACTOR RECEPTORS AND OTHER SIGNALING

CC PROTEINS THROUGH PROTEIN-TYROSINE PHOSPHORYLATION. THE SH2 REGIONS
 CC MAY INTERACT WITH OTHER CELLULAR COMPONENTS TO MODULATE ITS OWN
 CC PHOSPHATASE ACTIVITY AGAINST INTERACTING SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE)
 CC AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEMATOPOIETIC
 CC CELLS.
 CC -1- PFM: PHOSPHORYLATED ON SERINE AND TYROSINE.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. SUBCLASS THAT CONTAINS SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M74903; AAA35963.1; -;
 DR EMBL: X62053; CAA43982.1; -;
 DR EMBL: M77273; AAA3610.1; -;
 DR EMBL: U15528; AAA82880.1; -;
 DR EMBL: U15536; AAA82880.1; JOINED.
 DR EMBL: U15535; AAA82880.1; JOINED.
 DR EMBL: U15534; AAA82880.1; JOINED.
 DR EMBL: U15533; AAA82880.1; JOINED.
 DR EMBL: U15532; AAA82880.1; JOINED.
 DR EMBL: U15531; AAA82880.1; JOINED.
 DR EMBL: U15530; AAA82880.1; JOINED.
 DR EMBL: U15529; AAA82880.1; JOINED.
 DR EMBL: U15528; AAA82879.1; -;
 DR EMBL: U15537; AAA82879.1; JOINED.
 DR EMBL: U15535; AAA82879.1; JOINED.
 DR EMBL: U15534; AAA82879.1; JOINED.
 DR EMBL: U15533; AAA82879.1; JOINED.
 DR EMBL: U15532; AAA82879.1; JOINED.
 DR EMBL: U15531; AAA82879.1; JOINED.
 DR EMBL: U15530; AAA82879.1; JOINED.
 DR EMBL: U15529; AAA82879.1; JOINED.
 DR EMBL: U47924; AAB51322.1; -;
 DR EMBL: S20825; S20825.
 DR PIR: S20825; S20825.
 DR PIR: B42031; B42031.
 DR PDB: 1GWZ; 22-AUG-99.
 DR PDB: IFPR; 07-MAR-01.
 DR Genew; HGNC:9658; PTPN6.
 DR MIM: 176883; -;
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR InterPro: IPR000242; TYR_PP.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00102; Y_phosphatase; 1.
 DR PRINTS: PRO0700; PRYPPHATASE.
 DR PRINTS: PRO0401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2; 2.
 DR SMART: SM00194; PTPc; 1.
 DR SMART: SM00252; SH2; 2.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
 DR PROSITE: PS50001; SH2; 2.
 KW Hydrolase; SH2 domain; Repeat; phosphorylation; Alternative splicing;
 KW 3D-structure.
 FT DOMAIN 4 100 SH2 1.
 FT DOMAIN 110 213 SH2 2.
 FT DOMAIN 269 514 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 453 453

FT VARSPLIC 1 39 MISSING (IN SHORT ISOFORM).
 FT VARSPLIC 40 44 SLAYR -> MLSRG (IN SHORT ISOFORM).
 FT CONFLICT 1 3 MYR -> MLSRG (IN REF. 2).
 FT CONFLICT 6 6 H -> L (IN REF. 5).
 FT CONFLICT 86 86 L -> V (IN REF. 4).
 FT CONFLICT 146 146 V -> E (IN REF. 5).
 SO SEQUENCE 595 AA; 67561 MW; 4D736C21D3542D2 CMC64;
 Query Match 13.0%; Score 104; DB 1; Length 595;
 Best Local Similarity 27.5%; Pred. No. 0.0074;
 Matches 39; Conservative 17; Mismatches 36; Indels 50; Gaps 7;
 QY 27 HYQTLIDGVRLHSLVTERGPHSDCPGLTLRLRLDPFCPPADQIDRFVQYVDANA 86
 DB 411 HYQYL-----SMPDHGVP---SEPGVYL-----SFLDQINQ 438
 QY 87 RGEA-----GVGHALFGFRGTMLACYKE-----RGLAG-----DAIAEIRLRGCS 132
 DB 439 RQESLPHAGPLIVHCASIGRTGITIVDMLENTISTGLDCCDIDIOKTQMVVRAQRSGM 498
 QY 133 IETYEQER-----AVPQFYQRTK 150
 DB 499 VQTEHQYKFIYVAIAQFIETTK 520
 RESULT 15
 PTPN6_MOUSE STANDARD; PRT: 595 AA.
 ID PTPN6_MOUSE 063872; 063873; 063874; 09QVA6; 09QVA7; 09QVA8; 09ROV6;
 AC P29351; 063872; 063873; 063874; 09QVA6; 09QVA7; 09QVA8; 09ROV6;
 AC Q921G3; 035128;
 DT 01-DEC-1992 (Ref. 24, Created)
 DT 15-JUN-2002 (Ref. 41, Last sequence update)
 DT 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Protein-tyrosine phosphatase, non-receptor type 6 (EC 3.1.3.48)
 DE (Protein-tyrosine phosphatase 1C) (PTP-1C) (Hematopoietic cell
 DE protein-tyrosine phosphatase) (70Z-SHP) (SH-PTP1).
 GN PTPN6 OR PTP1C OR HCP OR HCPH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DKA/2;
 RX MEDLINE=92123209; PubMed=1732748;
 RA Yi T., Cleveland J.L., Ihle J.N.;
 RT "Protein tyrosine phosphatase containing SH2 domains:
 RT characterization, preferential expression in hematopoietic cells, and
 RT localization to human chromosome 12p12-p13.";
 RL Mol. Cell. Biol. 12:836-846(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=92236615; PubMed=1373816;
 RA Matthews R.J., Howe D.B., Flores E., Thomas M.L.;
 RT "Characterization of hematopoietic intracellular protein tyrosine
 RT phosphatases: description of a phosphatase containing an SH2 domain
 RT and another enriched in proline-, glutamic acid-, serine-, and
 RT threonine-rich sequences.";
 RL Mol. Cell. Biol. 12:2396-2405(1992).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANTS MOTHATEN AND VIABLE MOTHATEN.
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
 RX MEDLINE=9331972; PubMed=8324828;
 RA Schultz L.D., Schweitzer P.A., Rejan T.V., Yi T., Ihle J.N.,
 RA Matthews R.J., Thomas M.L., Beier D.R.;
 RT "Mutations at the murine motheaten locus are within the hematopoietic
 RT cell protein-tyrosine phosphatase (Hcpn) gene.";
 RL Cell 73:1445-1454(1993).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RX MEDLINE=98112780; PubMed=9445485;
 RA Ansari-Lari M.A., Celtyen J.C., Schwartz S., Zhang Z., Muzny D.M.,
 RA Lu J., Gorrell J.H., Chinnault A.C., Belmont J.W., Miller W.,

RA Gibbs R.A.;
 RT Comparative sequence analysis of a gene-rich cluster at human
 chromosome 12p13 and its syntenic region in mouse chromosome 6.;
 RL Genome Res. 8:29-40(1998).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND SUBUNIT.
 RC STRAIN-C3H; TISSUE-Adrenal gland;
 RX MEDLINE-9348302; PubMed-10419485;
 RA Martin A., Tsui H.W., Shulman M.J., Isenman D., Tsui F.W.;
 RT Murine SHP-1 splice variants with altered Src homology 2 (SH2)
 RT domains. Implications for the SH2-mediated intramolecular regulation
 of SHP-1.;
 RL J. Biol. Chem. 274:21725-21734(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RT Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 54-68, 128-135, 137-151; 242-252; 278-285; 293-308 AND
 RP 373-382, AND PHOSPHORYLATION.
 RX MEDLINE-93054686; PubMed-1385421;
 RA Yeung Y.-G., Berg K.L., Pixley F.J., Angeletti R.H., Stanley E.R.;
 RT Protein tyrosine phosphatase-1C is rapidly phosphorylated in tyrosine
 RT in macrophages in response to colony stimulating factor-1.;
 RL J. Biol. Chem. 267:23447-23450(1992).
 CC -1- FUNCTION: PLAYS A KEY ROLE IN HEMATOPOIESIS. THIS PTASE ACTIVITY
 CC MAY DIRECTLY LINK GROWTH FACTOR RECEPTORS AND OTHER SIGNALING
 CC PROTEINS THROUGH PROTEIN-TYROSINE PHOSPHORYLATION. THE SH2 REGIONS
 CC MAY INTERACT WITH OTHER CELLULAR COMPONENTS TO MODULATE ITS OWN
 CC PHOSPHATASE ACTIVITY AGAINST INTERACTING SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SUBUNIT: Monomer.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEMATOPOIETIC
 CC CELLS
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are
 CC produced by alternative splicing.
 CC -1- PTM: Phosphorylated on tyrosine residues.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. SUBCLASS THAT CONTAINS SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M68902; AAA37796.1; -;
 DR EMBL; M90389; AAA40007.1; -;
 DR EMBL; S63763; -; NOT_ANNOTATED_CDS.
 DR EMBL; S63764; -; NOT_ANNOTATED_CDS.
 DR EMBL; S63803; -; NOT_ANNOTATED_CDS.
 DR EMBL; AC002397; AAC36009.1; -;
 DR EMBL; AC002397; AAC36008.1; -;
 DR EMBL; U65955; AAD00152.1; -;
 DR EMBL; U65952; AAD00152.1; -;
 DR EMBL; U65953; AAD00152.1; JOINED.
 DR EMBL; U65954; AAD00152.1; JOINED.
 DR EMBL; U65955; AAD00151.1; -;
 DR EMBL; U65951; AAD00151.1; JOINED.
 DR EMBL; U65952; AAD00151.1; JOINED.
 DR EMBL; U65953; AAD00151.1; JOINED.
 DR EMBL; U65954; AAD00151.1; JOINED.
 DR EMBL; BC012660; AAH12660.1; -;
 DR PIR; A44390; A44390.
 DR HSSP; P29350; 1GW2.
 DR MGD; MGI:96055; Hcph.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR000387; TYR_phosphatase.

DR InterPro; IPR000242; TYR_PP.
 DR Pfam; PF00017; SH2; 2.
 DR Pfam; PF00102; Y-phosphatase; 1.
 DR PRINTS; PR00700; PRITPHPTASE.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR ProDom; PD000093; SH2; 2.
 DR SMART; SM00194; PTPC; 1.
 DR SMART; SM00252; SH2; 2.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS0055; TYR_PHOSPHATASE_PTP; 1.
 DR PROSITE; PS0001; SH2; 2.
 KW Hydrolyase; SH2 domain; Repeat; Phosphorylation; Alternative splicing.
 FT DOMAIN 4 100
 FT DOMAIN 110 213
 FT DOMAIN 269 514
 FT ACT_SITE 453 453
 FT VARSPPLIC 1 3
 FT VARSPPLIC 1 39
 FT VARSPPLIC 40 44
 FT VARIANT 77 99
 FT VARIANT 100 595
 FT CONFLICT 240 240
 FT CONFLICT 572 572
 FT CONFLICT 586 586
 SQ SEQUENCE 595 AA; 67559 MW; CFI7300D032638D2 CRC64;
 Query Match 13.0%; Score 104; DB 1; Length 595;
 Best Local Similarity 27.5%; Pred. No. 0.0074;
 Matches 39; Conservative 17; Mismatches 36; Indels 50; Gaps 7;
 QY 27 HYQFLDLGVRHLYSLTERPHSDSCPGTLRLAIIPFCPPAPQIDRFQIVDEANA 86
 DB 411 HYQYL-----SWPDHGV-----SEPGVL-----SFLDINO 438
 QY 87 RGEA-----VGVCALGFGRTGTMACYLKE-----RGLAAG-----DAIAEIRRLRPGS 132
 DB 439 RQSLRHAGRIIVHCAGIGRTITIVIMKESISTKGLDCDIDIKTIQWRAORSQM 498
 QY 133 IETYEQEK-----AVFOFYQRTK 150
 DB 499 VQTEAQYKEIYVAIAQFIETTK 520

Search completed: November 4, 2002, 12:47:03
 Job time : 27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 12:43:42 ; Search time 43 Seconds
(without alignments)
335.353 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798

Sequence: 1 MGVPNFSWVLPGRAGLA.....GSIETYEOKKAVFQYQRTK 150

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PTR_73:*

1: ptr1:*

2: ptr2:*

3: ptr3:*

4: ptr4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	242.5	30.4	190	2 T31661	hypothetical prote
2	186.5	23.4	681	2 E88158	protein C17G10.4c
3	186.5	23.4	708	2 T34098	hypothetical prote
4	186.5	23.4	1063	2 T34097	probable protein-t
5	157.5	19.7	161	2 B90417	conserved hypotet
6	152.5	19.1	151	2 G75157	protein tyrosine p
7	146	18.3	146	2 G71181	hypothetical prote
8	145.5	18.2	537	2 T50099	probable protein-t
9	143	17.9	446	2 T33986	hypothetical prote
10	142.5	17.9	383	2 F84771	hypothetical prote
11	137.5	17.2	551	2 S56283	protein-tyrosine-p
12	133	16.7	177	2 D75309	protein-tyrosine-p
13	125	15.7	376	2 S19740	hypothetical prote
14	119	14.9	168	2 T41891	phosphotyrosine ph
15	117	14.7	165	2 A82246	probable phosphat
16	115	14.4	437	2 DB3329	conserved hypotet
17	115	14.4	541	2 JC5936	mRNA guanylyltrans
18	115	14.4	597	2 JC5936	mRNA guanylyltrans
19	114	14.3	168	1 A40781	dual specificity p
20	114	14.3	398	2 T08716	protein-tyrosine-p
21	113	14.2	261	2 T25206	hypothetical prote
22	113	14.2	511	2 JC7198	caping enzyme 1b
23	113	14.2	585	2 T18885	hypothetical prote
24	113	14.2	598	2 JC7197	caping enzyme 1a
25	110	13.8	535	2 A46101	protein-tyrosine-p
26	110	13.8	548	2 B46101	protein-tyrosine-p
27	107	13.4	220	2 T10279	protein-tyrosine-p
28	105	13.0	177	2 A12400	hypothetical prote
29	104	13.0	276	2 T48906	protein-tyrosine-p

30	104	13.0	595	1 S20825	protein-tyrosine-p
31	104	13.0	595	1 A44390	protein-tyrosine-p
32	103.5	13.0	159	2 H64326	hypothetical prote
33	99	12.4	190	2 T43172	probable protein-t
34	98.5	12.3	699	2 JC6132	protein-tyrosine-p
35	98	12.3	360	1 JH0692	protein-tyrosine-p
36	98	12.3	700	1 S12053	protein-tyrosine-p
37	98	12.3	2051	2 T30938	receptor tyrosine
38	97.5	12.2	549	2 S53427	protein-tyrosine-p
39	97.5	12.2	656	1 A55574	protein-tyrosine-p
40	97.5	12.2	656	1 UC4263	protein-tyrosine-p
41	96.5	12.1	147	2 C64512	hypothetical prote
42	96.5	12.1	150	2 T21489	hypothetical prote
43	94.5	11.8	173	2 JC5982	prenylated protein
44	94.5	11.8	377	1 A48711	protein-tyrosine-p
45	94	11.8	284	2 T23421	hypothetical prote

ALIGNMENTS

RESULT 1

T31661 hypothetical protein COS41.7 - sea squirt (Cliona intestinalis)

C/Species: Clona intestinalis

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T31661

R/Blrd, A.P.; Clark, V.; Jones, S.J.; Lettgeb, S.; Dobson, R.; Tweedle, S.

submitted to the EMBL Data Library, December 1996

A/Reference number: Z21049

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-190 <BIR>

A/Cross-references: EMBL:283760; NID:e1014349; PID:e289997; PIDN:CAB06051.1

C/Genetics:

A/Introns: 56/1; 87/1

Query Match

Best Local Similarity 42.7%; Pred. No. 1.5e-16;

Matches 61; Conservative 18; Mismatches 39; Indels 25; Gaps 3;

QY	5	PPNFSWVLPGRAGLAALPRPAHYQFLDGVHIVSLFGRGPHSDSCPLTLRLRIP	64
DB	66	PSNFTVVDGKMLAACASPTCPRTYGQ-----VHSHSTLFR-----NTDAPFKL-----	107
QY	65	DECPAPDQIDRFVQIVDEANANGAEVGVHCAIGFRTGTMLACYLKERGLAAGDAIAE	124
DB	108	-----NOTVWFTKSNFKGFVYQAVAVHACANGKISGTMLACYLVKMKRISADALKE	160
QY	125	IRRLRGSIETYEOKKAVFQYQ	147
DB	161	IRMRPGSVSETEOKKAVFQYQ	183

RESULT 2

E88158 protein C17G10.4c [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C/Accession: E88158

R/Anonymous, The C. elegans Sequencing Consortium.

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating blo

A/Reference number: A55000; MUID:99069613; PMID:9851916

A/Note: see websites genstl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C-

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;

A/Accession: E88158

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-681 <STO>

A/Cross-references: GB:chr_II; PIDN:AB93458.1; PID:g7731379; GSPDB:GN00020; CESP:C17G

C/Genetics:

```

A:Gene: C17G10.4c
A:Map position: 2

Query Match          23.4%, Score 186.5; DB 2; Length 681;
Best Local Similarity 29.8%; Pred. No. 2.2e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

QY 3 VQPPNFSWTLPERLAGLALPR-----LPAN-----YQFLDLGVRLVSL----- 42
DB 192 VENGDFNMIIPEKILISFCGPHNRSRENGPYPHADVYFDFERENKYSTIYVLNKNYDA 251
QY 43 ---TERGPPHSDSCGLTLHLRLRIDPCFPAPADQIDRFQYVDENANGAIVGHCALGF 99
DB 252 SKFTAGAGDHD-----LFFIDGSTPDSDEIMLKFIKIVD--NTKG-GVAIVCKKAGL 299
QY 100 GRTGTMACIYLVKERGLAAGDAIAEIRLRIPGSI---ETYEQEKAVF 143
DB 300 GRTGTLICMWMKKEYGLTAGECMGMWRVCRCPSVIGPOOPYLIEKOKF 347

RESULT 3
T34098
hypoetical protein C17G10.4a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34098
R:Johnson, D.
Submitted to: the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid C17G10.
A:Reference number: Z21476
A:Accession: T34098
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-708 <JOB>
A:Cross-references: EMBL:U28739; PIDs: AAB93457.1; GSPDB: GN00020; CESP: C17G10.4a
A:Experimental source: strain Bristol N2; clone C17G10
C:Genetics:
A:Gene: CESP:C17G10.4a
A:Map position: 2
A:Introns: 19/3; 84/3; 120/3; 541/3; 605/3; 686/3

Query Match          23.4%, Score 186.5; DB 2; Length 708;
Best Local Similarity 29.8%; Pred. No. 2.3e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

QY 3 VQPPNFSWTLPERLAGLALPR-----LPAN-----YQFLDLGVRLVSL----- 42
DB 192 VENGDFNMIIPEKILISFCGPHNRSRENGPYPHADVYFDFERENKYSTIYVLNKNYDA 251
QY 43 ---TERGPPHSDSCGLTLHLRLRIDPCFPAPADQIDRFQYVDENANGAIVGHCALGF 99
DB 252 SKFTAGAGDHD-----LFFIDGSTPDSDEIMLKFIKIVD--NTKG-GVAIVCKKAGL 299
QY 100 GRTGTMACIYLVKERGLAAGDAIAEIRLRIPGSI---ETYEQEKAVF 143
DB 300 GRTGTLICMWMKKEYGLTAGECMGMWRVCRCPSVIGPOOPYLIEKOKF 347

RESULT 4
T34097
probable protein-tyrosine-phosphatase (EC 3.1.3.48) CDC14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000
C:Accession: T34097; T34099; T37328
R:Johnson, D.
Submitted to: the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid C17G10.
A:Reference number: Z21476
A:Accession: T34097
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1063 <JOB>
A:Cross-references: EMBL:U28739; PIDs: AAB93459.1; GSPDB: GN00020; CESP: C17G10.4b

```

[illegible]

RESULT 6

C75157 protein tyrosine phosphatase PAB0279 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: C75157

R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure

A:Reference number: A75001

A:Accession: C75157

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-151 <KAW>

A:Cross-references: GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB49338.1; PID:ej51532

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB0279

Query Match

Best Local Similarity 30.4%; Pred. No. 8.7e-08;

Matches 31; Conservative 27; Mismatches 35; Indels 9; Gaps 1;

OY 58 LHRLR-----IDFCPPADQIDRFVQIYDEANANGAAGVHCALGFGRTGTLAC 108

DB 43 LKRLRQAEVLHSPIDFAPSLQELMEIEMWEKREKRYIHICGSGSGTATA 102

OY 109 YLKERGLAAGDAIAEIRLRPGSIEFYDEKAVFOFYQTK 150

DB 103 WLMYSOGIPLREALRRRLKPSAVETEDQKILIEFEFEPLK 144

RESULT 7

G71181 hypothetical protein PH1732 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C:Accession: G71181

R:Kawabuchi, Y.; Sawada, M.; Horikawa, H.; Hatakeyama, Y.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushiida, N.; Oguchi, M. Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon

A:Reference number: A71000; MID:98344137; PMID:9679194

A:Accession: G71181

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-146 <KAW>

A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA30846.1; PID:g3258163

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH1732

C:Superfamily: VHL-type dual specificity phosphoprotein phosphatase homology F30-143/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

Query Match

Best Local Similarity 32.5%; Pred. No. 3.7e-07;

Matches 27; Conservative 24; Mismatches 32; Indels 0; Gaps 0;

OY 63 IDFCPPADQIDRFVQIYDEANANGAAGVHCALGFGRTGTLACIYKRGIAAGDAI 122

DB 57 IDFTAPSLQSLKYIKWEKREKRYIHICGSGSGTAVAVAMWISQGLSREGI 116

OY 123 AEIRLRPGSIEFYDEKAVFOF 145

DB 117 RRVRLKPSAVETEDQLEVLREF 139

Db 117 RRVRLKPSAVETEDQLEVLREF 139

RESULT 8

750099 Probable protein-tyrosine phosphatase CDC14 homolog [Imported] - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000

C:Accession: T50099

R:Mod. V.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.

submitted to the EMBL Data Library, February 2000

A:Reference number: Z25037

A:Accession: T50099

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-537 <MOO>

A:Cross-references: EMBL:AL158056; PIDN:CAB76271.1; GSPDB:GN00066; SPDB:SPAC1782.09c

A:Experimental source: strain 972h(-); cosmid c1782

C:Genetics:

A:Gene: SPDB:SPAC1782.09c

A:Map position: 1

A:Introns: 17/1

Query Match

Best Local Similarity 24.6%; Pred. No. 1.9e-06;

Matches 44; Conservative 28; Mismatches 72; Indels 35; Gaps 6;

OY 3 YQPNNSWVLPGLAGLALP-----RLPAHYOFLIDGLVRLVSLTER--GP 47

DB 178 VENGDENMWISP-KFIAFSPIDQAGMNRASRPKRLQPFPAIVIDYFVANKVLIVRLNGP 236

OY 48 PHSDSC---PGLTLRLRIPDPCPPADQIDRFVQIYDEANANGAAGVHCALGFGRTGT 104

DB 237 LYKRTFFENVGIRHKEMTFEDGVPELSYKFTIDLEEVEEDG-VIAVCKKAGLGRTGC 295

OY 105 MLCYLVKERGLAAGDAIAEIRLRPGSI-----EYDEKAVFOFYQTK 148

DB 296 LIGAVILYKHCFTANEVIAVYRIMRPGMVGPOQHMLHINQVHRAVYFKEMKAGRAIQ 354

RESULT 9

T33986 hypothetical protein T12B3.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33986

R:Nelson, J.; Wohlmann, P.; Antoniou, B.; Fulton, B.

submitted to the EMBL Data Library, February 1999

A:Description: The sequence of C. elegans cosmid T12B3.

A:Reference number: Z21450

A:Accession: T33986

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-446 <NEU>

A:Cross-references: EMBL:AF125960; PIDN:AD14737.1; GSPDB:GN00022; CESP:T12B3.1

A:Experimental source: strain Bristol N2; clone T12B3

C:Genetics:

A:Gene: CESP:T12B3.1

A:Map position: 4

A:Introns: 21/1; 100/2; 139/1; 228/3; 271/3; 351/3; 391/3

Query Match

Best Local Similarity 17.9%; Score 145; DB 2; Length 446;

Matches 37; Conservative 35; Mismatches 54; Indels 24; Gaps 3;

OY 19 LALPRLPAHYOFLIDP-----GVRHLVSLTERG-----PPHSDSCPG 55

DB 79 LAMARLOVEHFDLGIYEKFKTGIGSVINLQSGSHSCGSGNLISGFYDPEINLRNG 138

OY 56 LTIHLRLIPDPCPPADQIDRFVQIYDEANANGAAGVHCALGFGRTGTLACIYKRG 115

DB 139 ITHYNPLPDPFACFPNRLDLYKVVDFALSHK-IAYVCHAGHGRGTGVIAMAMVATLG 197

OY 116 IAAGDAIAEIRLRPGSIEFYDEKAVFOF 145

DB 198 MSPSQAVDTVRSRRAKAVOSKEQVKTILHEF 227

RESULT 10

F84771
 hypothetical protein At2g35680 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84771
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shua, T.P.; Bentos, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, I.A.; Shen, M.; Venkatesh, S.E.; Umayam, L.; Talion, L.;
 Neus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MID:20083487; PMID:10617197
 A:Accession: F84771
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-383 <STO>
 A:Cross-references: GB:AE002093; NID:94263787; PIDN:AA015447.1; GSPDB:GM00139
 C:Genetics:
 A:Gene: At2g35680
 A:Map position: 2

Query Match 17.9%; Score 142.5; DB 2; Length 383;
 Best Local Similarity 28.8%; Pred. No. 2.5e-06;
 Matches 42; Conservative 31; Mismatches 58; Indels 15; Gaps 6;

QY 10 WVLGRAGALPRRLPAHYFLDLGYRHLVSLFERGPPHSDCSGLTLR-----L 61
 DB 122 WFLQFLLG-AVP-FPSDVPQLKELGVCVITLNE--PPTLVPS-SLKYSCIDHLVI 175
 QY 62 RPPDFC-PPAPDQIDRFVQIVDEANRGEAVGYCALGFRGTGLMCIYKRGGLAAGD 120
 DB 176 ATMDYCAPMEALICQAVEEIHNRASLGKTTYHCKAGRGSRSTIVICIVLQHKMPEA 235
 QY 121 AIAEIRLRPGSTETYEGRKAFQFY 146
 DB 236 AYSYRSIRPRVLLAAQKRAVEY 261

RESULT 11
 S56283
 protein-tyrosine-phosphatase (EC 3.1.3.48) CDC14 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: cell division control protein CDC14; protein R005; protein YFR028C
 C:Species: Saccharomyces cerevisiae
 C:Date: 02-Sep-1995 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
 C:Accession: S56283; A42784; S38416; S59694; S62239; S63833; S69234
 R:Murkani, Y.; Naitou, M.; Hagihara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanu
 submitted to the EMBL Data Library, May 1995
 A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce
 A:Reference number: S56186
 A:Accession: S56283
 A:Molecule type: DNA
 A:Residues: 1-351 <MUR>
 A:Cross-references: EMBL:D50617; NID:9836685; PID:d1009908; PID:g836783; MIPS:YFR028C
 R:Man, J.; Xu, H.; Grunstein, M.
 J. Biol. Chem. 267, 11274-11280, 1992
 A:Title: CDC14 of Saccharomyces cerevisiae. Cloning, sequence analysis, and transcriptio
 A:Reference number: A42784; MID:92283835; PMID:1597462
 A:Accession: A42784
 A:Molecule type: DNA
 A:Residues: 'MPYI', 114, 'MRISQFQKM', 124-125, 'A', 127, 'G', 131, 'LRKRL', 139, 'C', 141, 'PSI'
 A:Cross-references: EMBL:M61194; NID:g171182; PID:g171183; EMBL:S37205
 A:Note: sequence extracted from NCBI backbone (NCBIN:104751, NCBIP:104754)
 R:Mal, B.; Lipp, M.
 submitted to the EMBL Data Library, August 1993
 A:Reference number: S38416
 A:Accession: S38416
 A:Molecule type: DNA
 A:Residues: 'M', 101, 'AL', 104, 'LYD', 110, 'RSGMDSASASAV', 118-421 <MAL>
 A:Cross-references: EMBL:X75077; NID:g407517; PID:g407518
 R:Masaki, S.M.S.; Yasushi, M.Y.M.; Akio, T.E.A.T.
 submitted to the EMBL Data Library, June 1995
 A:Description: Dominant mutants alleles of a yeast protein kinase gene CDC15 suppress th
 A:Reference number: S59694
 A:Accession: S59694

A:Molecule type: DNA
 A:Residues: 1-117, 'P', 119-551 <MAS>
 A:Cross-references: EMBL:D55715; NID:g870755; PID:d1010176; PID:g870756
 R:Murkani, Y.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S62230
 A:Accession: S62239
 A:Molecule type: DNA
 A:Residues: 1-551 <MUR>
 A:Cross-references: EMBL:PD4602; NID:g893419; PID:d1008630; PID:g893423
 R:Exl, T.; Naitou, M.; Hagihara, H.; Abe, M.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.
 Yeast 12, 177-190, 1996
 A:Title: Fifteen open reading frames in a 30.8 kb region of the right arm of chromoso
 A:Reference number: S63830; MID:96287654; PMID:8686381
 A:Accession: S63833
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1551 <EKT>
 A:Cross-references: EMBL:D44602; NID:g893419; PID:d1008630; PID:g893423
 R:Shirayama, M.; Matsui, Y.; Toh-e, A.
 Mol. Gen. Genet. 251, 176-185, 1996
 A:Title: Dominant mutant alleles of yeast protein kinase gene CDC15 suppress the Itel
 A:Reference number: S69234; MID:96242150; PMID:8668128
 A:Accession: S69234
 A:Molecule type: DNA
 A:Residues: 1-117, 'P', 119-551 <SHI>
 A:Cross-references: EMBL:D5715; NID:g870755; PIDN:BA09533.1; PID:g870756
 C:Genetics:
 A:Gene: SGD:CDK14
 A:Cross-references: SGD:S0001924; MIPS:YFR028C
 A:Map position: 6R
 A:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphat
 F:283/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:289/Binding site: substrate phosphate (Arg) #status predicted

Query Match 17.2%; Score 137.5; DB 2; Length 551;
 Best Local Similarity 28.0%; Pred. No. 1.2e-05;
 Matches 30; Conservative 20; Mismatches 34; Indels 23; Gaps 3;

QY 27 HYQFLDLGYRHLVSLFERGPPHSDCSGLTLRHLRIPDCPPAPDQIDRFVQIVDEANA 86
 DB 238 HFE---DIGIQLDLIEDG-----TCPDLSI-----VKNFVGAETIYK 274
 QY 87 RGEAVGYCALGFRGTGLMCIYKRGGLAAGDAIEIRLRGSI 133
 DB 275 RGGKIAVHCRAKAGRTGCLIGAHLYTYGFTANECIEFLPIRGWV 321

RESULT 12
 D75309
 protein-tyrosine phosphatase-related protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: D75309
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MID:20036896; PMID:10567266
 A:Accession: D75309
 A:Molecule type: DNA
 A:Residues: 1-117 <WHI>
 A:Status: preliminary
 A:Molecule type: DNA
 A:Cross-references: GB:AE000513; NID:g6459945; PIDN:AA011705.1; PID:g645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2161
 A:Map position: 1

Query Match 16.7%; Score 133; DB 2; Length 177;
 Best Local Similarity 38.2%; Pred. No. 8.8e-06;
 Matches 34; Conservative 15; Mismatches 32; Indels 8; Gaps 3;

QY 67 CFPAPDID---REVQIVDE---ANARGEAVGHGALGFGRTGTMACYLVKRGSLAAG 119
 DB 90 CFIYDQGVSDRARGFGLDELDTLDDGNVYVHCRGGLGRAGLTAACLLV--QAGMKPD 148
 QY 120 DAIAEIRRLRPGSIETFEQEKAVFOYQR 148
 DB 149 DAIALVTRKTRGAIENARQEQIFREFAEQ 177

RESULT 13

S19740
 hypothetical protein B - Rhodobacter capsulatus
 C:Species: Rhodobacter capsulatus
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 12-Feb-1999
 C:Accession: S19740
 R:Kelly, D.J.
 submitted to the EMBL Data Library, February 1992
 A:Reference number: S19739
 A:Accession: S19740
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-376 <KEI>
 A:Cross-references: EMBL:X63974
 C:Superfamily: Vhl-type dual specificity phosphoprotein phosphatase homology F.63-182/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>

Query Match 15.7%: Score 125; DB 2; Length 376;
 Best Local Similarity 30.4%; Pred. No. 0.00013;

Matches 38; Conservative 16; Mismatches 49; Indels 22; Gaps 3;

QY 31 LLDLGRHLVSLTE-----RQPHSDSCPGTLRLRLRIPFCPPAPDQIDR 76
 DB 55 IDMGAAHTLVTEPEDELGLKVPDLGTQGP-----AGMDMHLPLIADYSVTPAFEAR 109
 QY 77 EV---QIVDEANARGEAVGHGALGFGRTGTMACYLVKRGSLAAGDAIAEIRLRPGSI 133
 DB 110 WQAEGRVIAALRAGADVYVHCRGGLGRAGLTAACLLV--QAGMKPD 148
 QY 134 ETVYEQ 138
 DB 170 ETPAQ 174

RESULT 14

T41891
 phosphotyrosine phosphatase orf1 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
 C:Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV
 A:Variety: isolate T3
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
 C:Accession: T41891
 R:Gomi, S.; Majima, K.; Maeda, S.
 J. Gen. Virol. 80, 1333-1337, 1999
 A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
 A:Reference number: Z22020; MUID:95281911; PMID:10355780
 A:Accession: T41891
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-168 <KAM>
 A:Cross-references: EMBL:U33180; NID:g3745835; PIDN:AAC63820.1; PID:g3745973
 A:Experimental source: isolate T3
 C:Genetics:
 A:Note: ptp
 C:Superfamily: dual specificity phosphoprotein phosphatase DUSP3; Vhl-type dual specificity phosphoprotein phosphatase orf1 - Bombyx mori nuclear polyhedrosis virus (isolate T3)

Query Match 14.9%: Score 119; DB 2; Length 168;
 Best Local Similarity 32.6%; Pred. No. 0.0002;

Matches 29; Conservative 15; Mismatches 41; Indels 4; Gaps 2;

QY 55 GFLHLRLRIPFCPPAPDQIDREVQIVDEANAR--GEAVGHGALGFGRTGTMACYLVK 112
 DB 77 GLYKKIQVPGQTLPSSESTVOEFIDVETFEKCPGLVGVHCTGHNRTGIWVCRIAMH 136

QY 113 ERLGAAGDAIAEIRLRPGSIE--TYEOE 139
 DB 137 TLGIAPOEAINRPEKARGHKIERQNTYQD 165

RESULT 15

A82246
 probable phosphatase VC1070 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: A82246
 R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
 Chardson, D.; Ermolaeva, M.D.; Yamathayan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406853; PMID:10952301
 A:Accession: A82246
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-165 <HEI>
 A:Cross-references: GB:AE004188; GB:AE003852; NID:g9655530; PIDN:AAF94229.1; GSPDB:GN
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC1070
 A:Map position: 1

Query Match 14.7%: Score 117; DB 2; Length 165;
 Best Local Similarity 33.3%; Pred. No. 0.00031;

Matches 39; Conservative 15; Mismatches 43; Indels 20; Gaps 5;

QY 35 GVRHLVSLTEGCPHSDSCPGTLRLRLRIPFCPPAPDQIDRFVQ-----IVDEANAR 87
 DB 58 GVELELAIEYKA-----GLQWTFAPITEDDC--APDA--AFADQMCHSPALHQAISR 105
 QY 88 GEAVGHGALGFGRTGTMACYLVKRGSLAAGDAIAEIRLRPGSIETFEQEKAVFO 144
 DB 106 GEKVALLCHMGSGRTG--LLAAHLLLEKGMPLIESITVQVALRGAFTKEVQVGVHQ 161

Search completed: November 4, 2002, 12:49:51
 Job time : 45 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 12:43:02 ; Search time 82 Seconds
(without alignments)
376.915 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798
Sequence: 1 MGVPNPFNSWVLPGRLAGLALPRLPANHYOFLDLGVRHLVSLTERGSPPHSDSCPGTLTLR 60

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	100.0	150	4	Q9BVJ7
2	793	99.4	150	4	Q9NX48
3	725	90.9	144	11	Q9CWA8
4	242.5	30.4	190	5	P91585
5	186.5	23.4	681	5	P81300
6	186.5	23.4	709	5	P81299
7	186.5	23.4	1063	5	Q09976
8	170.5	21.4	447	4	Q91600
9	169	21.2	147	1	Q8X270
10	168.5	21.1	459	4	O43183
11	168.5	21.1	471	4	O60730
12	168.5	21.1	498	4	O60729
13	163.5	20.5	195	17	Q82X04
14	162.5	20.4	157	17	Q97472
15	160.5	20.1	383	4	O60728
16	160.5	20.1	594	4	Q9UNH5

17	160.5	20.1	623	4	O60727	O60727 homo sapien
18	157.5	19.7	151	17	Q97VZ7	Q97VZ7 sulfobolus
19	152.5	19.1	151	17	Q9V1L1	Q9V1L1 pyrococcus
20	151	18.9	553	5	Q9V1W7	Q9V1W7 drosophila
21	147.5	18.5	162	17	Q8U0C5	Q8U0C5 pyrococcus
22	147.5	18.5	580	4	O43171	O43171 homo sapien
23	146	18.3	146	17	O59385	O59385 pyrococcus
24	145.5	18.2	537	3	Q9P7H1	Q9P7H1 schizosacch
25	143	17.9	446	5	Q9UAX0	Q9UAX0 caenorhabdi
26	137.5	17.2	326	3	O05673	O05673 saccharomyc
27	137.5	17.2	551	3	O05180	O05180 saccharomyc
28	137	17.2	245	10	Q940U5	Q940U5 arabidopsis
29	137	17.2	337	10	Q9Z0P1	Q9Z0P1 arabidopsis
30	136	17.0	698	6	Q9N091	Q9N091 macaca fasc
31	136	17.0	698	4	Q8WX19	Q8WX19 homo sapien
32	133	16.7	177	16	Q9RS63	Q9RS63 delnoccocus
33	128.5	16.1	649	5	Q9VY44	Q9VY44 drosophila
34	128	16.0	362	10	Q9AUS9	Q9AUS9 oryza sativ
35	125	15.7	419	2	O07839	O07839 rhodobacter
36	124.5	15.6	542	3	Q9P8D4	Q9P8D4 candida alb
37	123.5	15.5	752	4	Q9ULE6	Q9ULE6 homo sapien
38	120	15.0	491	12	Q8QUP6	Q8QUP6 infectious
39	119	14.9	168	12	O92506	O92506 bombyx mori
40	117	14.7	165	16	Q9KT35	Q9KT35 vibrio chol
41	116	14.5	168	12	Q9DXA4	Q9DXA4 bombyx mori
42	115	14.4	437	16	Q910U5	Q910U5 pseudomonas
43	115	14.4	463	10	Q94JH9	Q94JH9 oryza sativ
44	115	14.4	581	11	Q9DC11	Q9DC11 mus musculu
45	115	14.4	597	4	Q8WUM8	Q8WUM8 homo sapien

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	150 AA.
Q9BVJ7			
AC	Q9BVJ7		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	Hypothetical 16.6 kDa protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=LUNG;		
RA	Strasbourg R.;		
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDJ databases.		
DR	EMBL: BC001140; AAH01140.1; -		
DR	InterPro: IPR000340; DS_phosphatase.		
DR	InterPro: IPR000387; TYR_phosphatase.		
DR	Pfam: PF00782; DSPC: 1.		
DR	PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.		
DR	PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.		
KW	Hydrolase; Hypothetical protein.		
SQ	SEQUENCE 150 AA; 1658 MW; 4B72EFA0434B1B5F CRC64;		
Query Match	100.0%; Score 798; DB 4; Length 150;		
Best Local Similarity	100.0%; Pred. No. 1.8e-71;		
Matches 150; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MGVPNPFNSWVLPGRLAGLALPRLPANHYOFLDLGVRHLVSLTERGSPPHSDSCPGTLTLR 60		
DB	1 MGVPNPFNSWVLPGRLAGLALPRLPANHYOFLDLGVRHLVSLTERGSPPHSDSCPGTLTLR 60		
QY	61 LRLPDCPPAPDQIDRFVQIVDEANARGAAGVGHGALGFRGTMTACTYLVKRGIAAD 120		
DB	61 LRLPDCPPAPDQIDRFVQIVDEANARGAAGVGHGALGFRGTMTACTYLVKRGIAAD 120		
QY	121 ATAERLRRLRPGSIETFEQEKAVGFQYQRTK 150		

```

Db 121 AIAEIRRLRGSIETYEOKAVFOFYORTK 150
|||||
RESULT 2
09CX48 PRELIMINARY; PRT; 150 AA.
ID 09CX48;
AC 09CX48;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CDNA FLJ20442 f1s, clone KAR04828.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isono T., Sugano S.,
RT "NEDO human cDNA sequencing project."
RL EMBL: AK000449; BAA91172.1;
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSpC; 1.
DR SMART: SM00012; PTPC_DSpC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 150 AA; 16598 MW; 4B72EDCD34B1B5F CRC64;

Query Match 99.4%; Score 793; DB 4; Length 150;
Best Local Similarity 99.3%; Pred. No. 5.7e-71;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVOPPNFSVNLGRLAGLPLRPAYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 60
DB 1 MGVOPPNFSVNLGRLAGLPLRPAYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 60
QY 61 LRIPDFCPAPDIDRFVQIVDEANRGEAVGVHCAFGFGTMTLACYLKERGLAAGD 120
DB 61 LRIPDFCPAPDIDRFVQIVDEANRGEAVGVHCAFGFGTMTLACYLKERGLAAGD 120
QY 121 AIAEIRRLRGSIETYEOKAVFOFYORTK 150
DB 121 AIAEIRRLRGSIETYEOKAVFOFYORTK 150

RESULT 3
09CW48 PRELIMINARY; PRT; 144 AA.
ID 09CW48;
AC 09CW48;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE 1300005N15R1k protein (Fragment).
GN 1300005N15R1k.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-C57BL/6J; TISSUE=LIVER;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawada J., Shingawa A., Shidara K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,

```

```

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaelets P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK004912; BAB23663.1;
DR MGI: MGI:1915690; 1300005N15R1k.
DR InterPro: IPR000387; TYR_phosphatase.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
FT NON_TER
SQ SEQUENCE 144 AA; 16043 MW; ED7BA64AABBD1052 CRC64;

Query Match 90.9%; Score 725; DB 11; Length 144;
Best Local Similarity 94.4%; Pred. No. 3e-64;
Matches 136; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 NFSVNLGRLAGLPLRPAYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 66
DB 1 NFSVNLGRLAGLPLRPAYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 66
QY 67 CPAPDIDRFVQIVDEANRGEAVGVHCAFGFGTMTLACYLKERGLAAGDAIEIR 126
DB 61 CPAPDIDRFVQIVDEANRGEAVGVHCAFGFGTMTLACYLKERGLAAGDAIEIR 120
QY 127 RLRGSIETYEOKAVFOFYORTK 150
DB 121 RLRGSIETYEOKAVFOFYORTK 144

RESULT 4
ID P91585 PRELIMINARY; PRT; 190 AA.
AC P91585;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE COS4.7.
OS Clona Intestinalis.
OC Eukaryota; Metazoa; Chordata; Trochordata; Ascidacea; Enterogona;
OC Phlebobranchia; Clonidae; Clona.
ON NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird A.P., Clark V., Jones S.J.M., Leitgeb S., Dobson R., Tweedie S.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98208558; PubMed=9539755;
RA Stimm M.W., Leitgeb S., Clark V.H., Jones S.J.M., Bird A.;
RT "Gene number in an invertebrate chordate, Clona intestinalis."
RL Proc. Natl. Acad. Sci. U.S.A. 95:4437-4440(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148102; PubMed=10024242;
RA Stimm M.W., Leitgeb S., Charlton J., Jones S.J.M., Harris B.R.,
RT "Nonmethylated transposable elements and methylated genes in a
RT chordate genome."
RL Science 283:1164-1167(1999).
RN [4]
RP SEQUENCE FROM N.A.
RA Stimm M.W., Bird A.;

```

Sequence analysis of transposable elements in the Sea Squirt, Ciona intestinalis.
 RT Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 283760; CAB06051.1; .

DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR SMART; SM00012; PTPC_DSPC; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 SQ SEQUENCE 190 AA; 21292 MW; F8934267DDAD27E CRC64;

Query Match 30.4%; Score 242.5; DB 5; Length 190;
 Best Local Similarity 42.7%; Pred. No. 2.8e-16;
 Matches 61; Conservative 18; Mismatches 39; Indels 25; Gaps 3;

OY 5 PNFNSVTLGRGLAGLAPRLPAHYOPLDLGVNHLVSLTERGPPHSDSCGLTLHLRLIP 64
 DB 66 PNFNSVTLGRGLAGLAPRLPAHYOPLDLGVNHLVSLTERGPPHSDSCGLTLHLRLIP 107
 OY 65 DECPAPDQIDRFVQIVDEANARGEAVGVCALFGRTGMLACYLKERGLAAGDAIAE 124
 DB 108 -----NCTVFTKSNFKEFGVYQAVNHCANGKISGTMACYLVKKRISAADALKE 160
 OY 125 IRLRPGSIETVEQEKAVFOFYQ 147
 DB 161 IRRMRPGSVETSEQEKAVFOFYQ 183

RESULT 5
 P81300 PRELIMINARY; PRT; 681 AA.

AC P81300;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Hypothetical 76.7 kDa protein C17G10.4C in chromosome II.
 GN C17G10.4 OR CDC14.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Johnson D.;
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RA Ernsing B.R., Li L., Wishart M.J., Dixon J.E.;
 RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
 CC -1 ALTERNATIVE PRODUCTS: THE ORFS C17G10.4A, C17G10.4B AND C17G10.4C
 CC REPRESENT PROBABLE ALTERNATIVE SPLICED FORMS OF A YET HYPOTHETICAL
 CC PROTEIN.

CC -1 SIMILARITY: SOME TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.

DR EMBL; U28739; AAB93458.1; .
 DR EMBL; AF000363; AAB94407.1; .
 DR WormPep; C17G10.4C; CE08288.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; TYR_PP.
 DR Pfam; PF00102; Y_phosphatase; 1.
 DR SMART; SM00012; PTPC_DSPC; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1;
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 KW Hypothetical protein; Alternative splicing.
 SQ SEQUENCE 681 AA; 76698 MW; 2C056CD6FA03ADDC CRC64;

Query Match 23.4%; Score 186.5; DB 5; Length 681;
 Best Local Similarity 29.8%; Pred. No. 4.3e-10;
 Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

OY 3 VQPPNFSWVLPGRGLALPR-----LPAA-----YQFLDLGVNHLVSL----- 42
 DB 192 VENGDFNMIIPKILSFQCPHNSRENGPYIAPDYIDFYENKYSTIVRLNAKNYDA 251
 OY 43 ---TERGPPHSDSCGLTLHLRLIPDPCPAPDQIDRFVQIVDEANARGEAVGVCALGF 99
 DB 252 SKFTKAGFDHVD-----LEFIDGSTPSDEIMKFKIVVD--NTKG-GVAVHCKAGL 299
 OY 100 GRGTMLACYLKERGLAAGDAIAEIRLRPGSI-----ETVEQEKAVF 143
 DB 300 GRGTMLACMMKKEYGLTAGECGMWLRVCRPGSVIGPOQPYLIEKOF 347

RESULT 6
 P81299 PRELIMINARY; PRT; 709 AA.

AC P81299;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Hypothetical 79.7 kDa protein C17G10.4A in chromosome II.
 GN C17G10.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Johnson D.;
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.

RN [2]
 RP REVISIONS.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 CC -1 ALTERNATIVE PRODUCTS: THE ORFS C17G10.4A, C17G10.4B AND C17G10.4C
 CC REPRESENT PROBABLE ALTERNATIVE SPLICED FORMS OF A YET HYPOTHETICAL
 CC PROTEIN.

CC -1 SIMILARITY: SOME TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
 DR EMBL; U28739; AAB93457.1; .
 DR WormPep; C17G10.4A; CE16860.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; TYR_PP.
 DR Pfam; PF00102; Y_phosphatase; 1.
 DR SMART; SM00012; PTPC_DSPC; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1;
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 KW Hypothetical protein; Alternative splicing.
 SQ SEQUENCE 709 AA; 79739 MW; ABS873D8CFEBC0C CRC64;

Query Match 23.4%; Score 186.5; DB 5; Length 709;
 Best Local Similarity 29.8%; Pred. No. 4.5e-10;
 Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

OY 3 VQPPNFSWVLPGRGLALPR-----LPAA-----YQFLDLGVNHLVSL----- 42
 DB 192 VENGDFNMIIPKILSFQCPHNSRENGPYIAPDYIDFYENKYSTIVRLNAKNYDA 251
 OY 43 ---TERGPPHSDSCGLTLHLRLIPDPCPAPDQIDRFVQIVDEANARGEAVGVCALGF 99
 DB 252 SKFTKAGFDHVD-----LEFIDGSTPSDEIMKFKIVVD--NTKG-GVAVHCKAGL 299
 OY 100 GRGTMLACYLKERGLAAGDAIAEIRLRPGSI-----ETVEQEKAVF 143
 DB 300 GRGTMLACMMKKEYGLTAGECGMWLRVCRPGSVIGPOQPYLIEKOF 347

```
RESULT 7
ID 009976 PRELIMINARY; PRT; 1063 AA.
AC 009976; 009955;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 120.2 kDa protein.
GN C17G10.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; Pubmed=9851916;
RA None;
RT Investigating biology. The C. elegans Sequencing Consortium.
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Johnson D.;
RT "The sequence of C. elegans cosmid C17G10."
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28739; AAB93459.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam; PF00102; Y_phosphatase; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase; Hypothetical protein.
SQ SEQUENCE 1063 AA; 120248 MW; 2B0494459F710197 CRC64;

Query Match 23.4%; Score 186.5; DB 5; Length 1063;
Best Local Similarity 29.8%; Pred. No. 7.2e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

QY 3 VQPPNFSWVLPGRLAGLALP-----LPAH-----YQFLDLDGVRHLVSL----- 42
DB 192 VENGDEFWIIPGKILSFSGPHNSRENGYYPYHAPDYFFYFRNKYSTIRLNAKNYDA 251
QY 43 ---TENGPPHSDSCPGITLRLRIPDPAPDQIDRFVOIVDEANRGEAVGVHCAIGF 99
DB 252 SKFTKAGFDHVD-----LFFIDGSTPSDEIMLKFIKIVD--NFKG-GVAVHCAAGL 299
QY 100 GRGTGMIACLYLVERGILAGDAIAELRLRPGSI---ETFEDEKAVF 143
DB 300 GRGTGLIACWMAKKEYGITAGECMGLRVCRRPGSVIGPQGYLLIKRKF 347

RESULT 8
ID 09Y600 PRELIMINARY; PRT; 447 AA.
AC 09Y600;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE WUGSC; H_DJ1166619.1 protein.
GN WUGSC; H_DJ1166619.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063792; Pubmed=9847074;
RA Sulston J.E., Waterston R.;
RT "Toward a complete human genome sequence."
RL Genome Res. 8:1097-1108(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Zidanic M., Wohldman P., Holmes A.;
RT "The sequence of Homo sapiens PAC clone RP5-1166619."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006024; AAD15415.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam; PF00102; Y_phosphatase; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 447 AA; 51614 MW; 962ACC30F16D3EF CRC64;

Query Match 21.4%; Score 170.5; DB 4; Length 447;
Best Local Similarity 29.4%; Pred. No. 1e-08;
Matches 50; Conservative 27; Mismatches 62; Indels 31; Gaps 7;

QY 7 NFSWVLPGRLAGLALP-----RLPAHY-----QFLDLDGVRHLVSLPFR--GPPHS 50
DB 185 DLNMIIRDFRIFGCPHSKARLESGYIOHSPETIYQFKNNVTTIRLKNMADARFT 244
QY 51 DSCPGITLRLRIPDPAPDQIDRFVOIVDEANRGEAVGVHCAIGFRTGTMACYL 110
DB 245 DA-GFDHDLDFADGSTPDAIVKRFLLDICE--NAGS-AIAYHCKAGLGTGTLIACYI 299
QY 111 VKENGIAAGDAIAELRLRPGSITFEDEKAVFO-----FYQRTK 150
DB 300 MKHYRMTAELTIAWVRICRPGLVIGPOQFLVMKQTSIMLEGDYFRORLK 349

RESULT 9
ID 08X270 PRELIMINARY; PRT; 147 AA.
AC 08X270;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Protein tyrosine phosphatase.
GN TK-PTP.
OS Pyrococcus kodakarensis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=69014;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KOD1;
RA Jeon S., Fujiwara S., Takagi M., Fukui K., Imanaka T.;
RT "TK-PTP, protein tyrosine phosphatase from hyperthermophilic archaeon Thermococcus kodakarensis KOD1: Enzymatic characteristics and isolation of its substrate proteins."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051898; BAB83049.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam; PF00782; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
```

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update),

Matches	46;	Conservative	27;	Mismatches	60;	Indels	21;	Gaps	6
---------	-----	--------------	-----	------------	-----	--------	-----	------	---

```

QY 7 NFSWVLEGRGLALP-----RLPAHY-----QFLDLGVNHLVSLTER---GPPHS 50
DB 215 DLNMIIPDRFAFCGSPSRALSGYHQHSPETRYQYFKNNVTTIRLNRMADAKRT 274
QY 51 DSCGTLRLRLRIPDPCPPAPDQIDRFVQVDEANRGEAVGVCALGFRTGTMACYL 110
DB 275 DA--GPHHDLFFADGSGTPDAIVKEFLDICE--NAEG-AIAYVCKAGLGTGTLIACY 329
QY 111 VKERGLAAGDAIARLRIPGSIETYEOKRAVQ 144
DB 330 MKHYMTAAETIAWVIRICRPGSVIGPQOQFLVMK 363

RESULT 13
Q82X04 PRELIMINARY; PRT; 195 AA.
AC 082X04;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Conserved protein (possible dual specificity phosphatase).
GN PAE1536.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL: AE009826; AAL6345.1; -.
DR InterPro: IPR000387; TYR_Phosphatase.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR Complete proteome.
SQ SEQUENCE 195 AA; 22170 MW; DE2648D5987A8D41 CRC64;

Query Match 20.5%; Score 163.5; DB 17; Length 195;
Best Local Similarity 34.0%; Pred. No. 1.9e-08;
Matches 50; Conservative 21; Mismatches 59; Indels 17; Gaps 5;

QY 10 WVLPGRLAGLALPRLPAHYQFLDLGVNHLVSLTERGPPHSDSCPGTLRLRIP 64
DB 40 WIRP-RIAGSCMRG-RQDIEKMAELGKTVYSLAEANEIETYGWGLLEKTKLMKGMK 97
QY 64 -----PDCPPAPDQIDRFVQVDEANRGEAVGVCALGFRTGTMACYLKERGLA 117
DB 98 WIHWPDPDGPFR--KIDELVEILKEAKAGSYV-VHCVGIGRTPTVLAAYLATKMK 154
QY 118 AGDAIAIRLRIPGSIETYEOKRAVQ 144
DB 155 ADDAIRVERVNPVSLTDQYVALL 181

RESULT 14
Q974T2 PRELIMINARY; PRT; 157 AA.
AC 0974T2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein ST0580.
GN ST0580.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Arai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7.";
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000383; BAB5573.1; -.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00102; Y_phosphatase; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PP; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 157 AA; 18569 MW; B77C0A14DA6B061 CRC64;

Query Match 20.4%; Score 162.5; DB 17; Length 157;
Best Local Similarity 33.8%; Pred. No. 1.9e-08;
Matches 49; Conservative 24; Mismatches 59; Indels 13; Gaps 5;

QY 10 WVLPGRLAGLALPRLPAHYQFLDLGVNHLVSLTERGPPHSDSCPG---LTIRLRIP 64
DB 4 WVRGIIGSPIPTDELDENKREGVKRILLLEEW--ELEWAGSMDYFSLIREKGF 61
QY 65 DFC-PPAPD-----QIDRFVQVDEANRGEAVGVCALGFRTGTMACYLKERGLA 119
DB 62 EFLHEPIPDGAPFPDQFLRIY-EMLKGYANLVHCVGIGRTGTITAGYLMEFDSDG 120
QY 120 DAIAIRLRIPGSIETYEOKRAVQ 144
DB 121 EATIEVRNRPAGAVQTEQQLFLQ 145

RESULT 15
Q60728 PRELIMINARY; PRT; 383 AA.
AC 060728;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Cdc14A3 phosphatase.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Hao L., Baskerville C., Charbonneau H.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF064103; AAC16660.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR00242; TYR_PP.
DR Pfam: PF00102; Y_phosphatase; 1.
DR SMART: SM00012; PRPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 383 AA; 43908 MW; 28295D04793D00B7 CRC64;

Query Match 20.1%; Score 160.5; DB 4; Length 383;
Best Local Similarity 29.9%; Pred. No. 8.4e-08;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

QY 3 VQPNFSWVLEGRGLA-----GLALPRLPAHYQFLDLGVNHLVSL----- 42

```

```

Db 175 VENGDEPMIYPGKFLAFSGPHPKSKIENGYPPLHAPFAYPFPYFKKHNTTAVVRLNKKIYEA 234
QY 43 ---TERGPPHSDSCPGTLTLRLRIPDCPPAPQIDRFQIVDEANARGAVGVHCALGF 99
Db 235 KRFTDAGFEHYD-----LFFIDGSTPSDNIVRRFLNICE--NTEG-AIAVHCKAGL 282
QY 100 GRTGTMACYLVKERGIAAGDAIAEIRLRPGSI 133
Db 283 GRTGTLIACYVMKHRYRTHAEIILAMIRICRPGSI 316

```

Search completed: November 4, 2002, 12:48:47
 Job time : 84 secs

FT Domain 74..125
FT /note= "Tyrosine specific protein phosphatase
FT domain signature"
FT Modified-site 41
FT /note= "O-phosphorylated"
FT Modified-site 43
FT /note= "O-phosphorylated"
FT Modified-site 7
FT /note= "N-glycosylated"
XX MO200120004-A2.
XX 22-MAR-2001.
XX 14-SEP-2000; 2000WO-US25515.
XX 15-SEP-1999; 99US-0154141.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,
PI Lu DM;
XX WPI: 2001-244811/25.
DR N-PSDB: AAF30482.
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis,
PT treatment and prevention of gastrointestinal, immune system,
PT neurological and cell proliferative disorders -
XX
XX Claim 1; Page 91; 103pp; English.
XX
XX The present sequence is that of novel human protein phosphatase and
CC kinase protein PPHK-7, as predicted from Incyte Clone ID NO.
CC 1606974CB1 (see AAF30482). Tissues that express PPHK-7 (as a
CC fraction of total tissues expressing PPHK-7) include reproductive
CC (0.372), cardiovascular (0.140) and endocrine (0.093). Diseases or
CC conditions associated with tissues expressing PPHK-7 (as a
CC fraction of total tissues expressing PPHK-7) include cancer
CC (0.558), inflammation or trauma (0.233) or cell proliferation
CC (0.209). The encoded protein shows homology to Clona intestinalis
CC tyrosine phosphatase. The invention provides human PPHK-1 to -11
CC polypeptides (see AAB20322-32) and polynucleotides (see AAF30476-86).
CC It also provides expression vectors, host cells, antibodies, agonists
CC and antagonists, as well as methods for diagnosing, treating or
CC preventing disorders associated with expression of PPHK, including
CC gastrointestinal disorders, immune system disorders, neurological
CC disorders and cell proliferative disorders, including cancer.
XX
SQ Sequence 150 AA;
Query Match 100.0%; Score 798; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.2e-85;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVPNPNFNSWVLPGRLAGLALPRLPAHYQFLDLGVRHLVSTERGPHSDSCPGLTLHR 60
DB 1 MGVPNPNFNSWVLPGRLAGLALPRLPAHYQFLDLGVRHLVSTERGPHSDSCPGLTLHR 60
QY 61 LRIPDFCPAPDQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
DB 61 LRIPDFCPAPDQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
QY 121 AIAEIRRLRPGSIETYEOKAVFOFYORTK 150
DB 121 AIAEIRRLRPGSIETYEOKAVFOFYORTK 150
RESULT 2
ID AAB35275 standard; Protein: 150 AA.
XX AAB35275;
AC

XX 08-MAY-2001 (first entry)
DT Human dual specificity phosphatase DSP-11.
XX
XX Human; dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
DE MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
XX KMF graft-versus-host disease; autoimmune disease; metabolic disease.
XX
XX Homo sapiens.
XX
XX MO200105983-A1.
XX 25-JAN-2001.
XX 19-JUL-2000; 2000WO-US19710.
XX 20-JUL-1999; 99US-0144557.
XX
XX (CEPT-) CEPTYR INC.
XX
XX LucHe RM, Wel B;
PI
XX WPI: 2001-147348/15.
DR N-PSDB: AAF27958.
XX
XX Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase. Is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
XX Claim 1; Fig 2; 65pp; English.
XX
XX The present invention provides the protein and coding sequences of the
CC human dual-specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signalling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC human DSP-11 protein.
XX
SQ Sequence 150 AA;
Query Match 100.0%; Score 798; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.2e-85;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVPNPNFNSWVLPGRLAGLALPRLPAHYQFLDLGVRHLVSTERGPHSDSCPGLTLHR 60
DB 1 MGVPNPNFNSWVLPGRLAGLALPRLPAHYQFLDLGVRHLVSTERGPHSDSCPGLTLHR 60
QY 61 LRIPDFCPAPDQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
DB 61 LRIPDFCPAPDQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
QY 121 AIAEIRRLRPGSIETYEOKAVFOFYORTK 150
DB 121 AIAEIRRLRPGSIETYEOKAVFOFYORTK 150
RESULT 3
ID ABB43078 standard; Protein: 179 AA.
XX ABB43078;
XX 22-AUG-2002 (first entry)
XX Human ovarian antigen HVCAJ27, SEQ ID NO:4210.
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW

KM inflammatory condition; immune disorder; blood disorder;
KM cardiovascular disorder; respiratory disorder; neurological disorder;
KM gastrointestinal disorder; urinary system disorder; drug screening;
KM gene therapy; chromosome mapping; forensic analysis;
KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KM antiinflammatory; gynaecological; reproductive.
XX
OS Homo sapiens.
XX
PN WO200200677-A1.
XX
PD 03-JAN-2002.
XX
PE 07-JUN-2001; 2001WO-US18569.
XX
PR 07-JUN-2000; 2000US-209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-147878/19.
XX
N-PSDB; ABQ56155.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX

Claim 11; SEQ ID NO 4210; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 179 AA:

Query Match 100.0%; Score 798; DB 23; Length 179;
Best Local Similarity 100.0%; Pred. No. 5.2e-85;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGVPNPFNSWVLPGRLAGLALPRLPAHYQFLDLDGVRHLVSLTERGPPHSDSCPGTLNR 60
DB 30 MGVPNPFNSWVLPGRLAGLALPRLPAHYQFLDLDGVRHLVSLTERGPPHSDSCPGTLNR 89
OY 61 LRIPDFCPAPPOIDRFVQIVDEANRGAAGVGHGALGFGRTGTMLACTLVKRGGLAAGD 120
DB 61 LRIPDFCPAPPOIDRFVQIVDEANRGAAGVGHGALGFGRTGTMLACTLVKRGGLAAGD 120

DB 90 LRIPDFCPAPPOIDRFVQIVDEANRGAAGVGHGALGFGRTGTMLACTLVKRGGLAAGD 149
OY 121 AIAEIRLRPGSTIETYEOKAVFOYORTK 150
DB 150 AIAEIRLRPGSTIETYEOKAVFOYORTK 179

RESULT 4
AAB73231
ID AAB73231 standard; Protein; 150 AA.
XX
AC AAB73231;
XX
DT 11-MAY-2001 (first entry)
XX
DE Human phosphatase BAA91172_h.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
XX cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
XX congenital muscle disorder; Papillon-Leleuve syndrome; Cowden disease;
XX Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
XX schizophrenia; hamartoma.
XX
OS Homo sapiens.
XX
PN WO200112819-A2.
XX
PD 22-FEB-2001.
XX
PE 11-AUG-2000; 2000WO-US22158.
XX
PR 13-AUG-1999; 99US-0149005.
XX
PA (SUG-) SUGEN INC.
XX

PI PLOWMAN GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX
DR WPI; 2001-211226/21.
XX
N-PSDB; AAF63583.
XX

PT New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders -
XX
PS Claim 6; Fig 5; 138pp; English.

XX The present invention relates to phosphatase proteins and coding
XX sequences. The present sequence is one such phosphatase. Phosphatases are
XX enzymes that catalyse the dephosphorylation of proteins modified by
XX phosphorylation of serine, threonine or tyrosine residues. The
XX phosphatases are useful for treating a variety of diseases: for example
XX cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,
XX synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
XX hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
XX glioblastoma, colorectal cancer and thyroid cancer; pathophysiological
XX hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
XX congenital muscle disorders, Papillon-Leleuve syndrome, Cowden disease,
XX ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
XX Zonana syndrome, schizophrenia and hamartomas.
XX

Sequence 150 AA:

Query Match 99.4%; Score 793; DB 22; Length 150;
Best Local Similarity 99.3%; Pred. No. 1.6e-84;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MGVPNPFNSWVLPGRLAGLALPRLPAHYQFLDLDGVRHLVSLTERGPPHSDSCPGTLNR 60
DB 1 MGVPNPFNSWVLPGRLAGLALPRLPAHYQFLDLDGVRHLVSLTERGPPHSDSCPGTLNR 60
OY 61 LRIPDFCPAPPOIDRFVQIVDEANRGAAGVGHGALGFGRTGTMLACTLVKRGGLAAGD 120
DB 61 LRIPDFCPAPPOIDRFVQIVDEANRGAAGVGHGALGFGRTGTMLACTLVKRGGLAAGD 120


```

RESULT 7
ABB07846
ID ABB07846 standard; protein; 150 AA.
XX
AC ABB07846;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human tyrosine phosphatase-like enzyme.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cyostatic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antiarrhythmic; hypotensive; vulnerrary; gene therapy; human.
XX
OS Homo sapiens.
XX
FH Key
FH Region
FH 92.108
FH /note="the encoding nucleotide fragment for the above
FH residues is not indicated in the corresponding
FH DNA sequence"
XX
XX WO200220747-A2.
XX
XX 14-MAR-2002.
XX
XX 05-SEP-2001; 2001WO-EP10205.
XX
XX 11-SEP-2000; 2000US-231568P.
XX
XX 06-DEC-2000; 2000US-251403P.
XX
XX (FARB ) BAYER AG.
XX
XX kossida S;
XX
XX WPI; 2002-339803/37.
XX
XX N-PSDB; ABL40806.
XX
XX New human tyrosine phosphatase-like enzyme polypeptide, regulators of
XX which are useful for preventing, treating diabetes, obesity, cancer,
XX cardiovascular and pulmonary diseases
XX
XX Claim 18; Fig 2; 117pp; English.
XX
XX The invention relates to a purified human tyrosine phosphatase-like
XX enzyme polypeptide. The enzyme can be expressed by standard recombinant
XX methodology. The tyrosine phosphatase-like enzyme and encoding
XX polynucleotides are useful for screening for modulators which are used
XX for treating a tyrosine phosphatase-like enzyme dysfunction related
XX disease such as a central nervous system (CNS) disorder, diabetes,
XX obesity, chronic obstructive pulmonary disease, cardiovascular disease,
XX cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
XX other eating disorder such as bulimia. CNS disorders include brain
XX injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
XX Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
XX Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
XX immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
XX and cardiovascular diseases include myocardial infarction, ischaemic
XX diseases of the heart, atrial and ventricular arrhythmia, hypertensive
XX vascular diseases and peripheral vascular diseases. The enzyme is useful
XX in diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to diseases or abnormalities related to the presence of
XX mutations in the encoding nucleic acid sequences. The present sequence
XX represents the human tyrosine phosphatase-like enzyme polypeptide.
XX
XX Sequence 150 AA:

```

```

Query Match 99.4%; Score 793; DB 23; Length 150;
Best Local Similarity 99.3%; Pred. No. 1,6e-84;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MGVPNPNFNSWVLPGRLAGLALPRLPAAHYQFLDGVHIVSLTERGPHSDSCGTLHR 60
DB 1 MGVPNPNFNSWVLPGRLAGLALPRLPAAHYQFLDGVHIVSLTERGPHSDSCGTLHR 60
QY 61 LRIPDFCPAPDDQIDRFVQIVDEANANGEAVGACAFGFGTGMACYLKKEGGLAAGD 120
DB 61 LRIPDFCPAPDDQIDRFVQIVDEANANGEAVGACAFGFGTGMACYLKKEGGLAAGD 120
QY 121 AIAEIRRLRPGSIETTYEKEAVQFYQRTK 150
DB 121 AIAEIRRLRPGSIETTYEKEAVQFYQRTK 150

RESULT 8
AAB73211
ID AAB73211 standard; protein; 150 AA.
XX
AC AAB73211;
XX
DT 11-MAY-2001 (first entry)
XX
DE Murine phosphatase AA023073_m.
XX
KW Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamartoma.
XX
XX Mus sp.
XX
XX WO200112819-A2.
XX
XX 22-FEB-2001.
XX
XX 11-AUG-2000; 2000WO-US22158.
XX
XX 13-AUG-1999; 99US-0149005.
XX
XX (SUGB-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Hall R, Flanagan P, Lloblin M;
XX
XX WPI: 2001-211226/21.
XX
XX N-PSDB; AAF63563.
XX
XX New protein phosphatase polypeptide for diagnosing and treating
XX phosphatase related disorders such as cancer, schizophrenia, cardiac
XX dysfunction and/or vascular disorders
XX
XX Claim 6; Fig 5; 138pp; English.
XX
XX The present invention relates to phosphatase proteins and coding
XX sequences. The present sequence is one such phosphatase. Phosphatases are
XX enzymes that catalyse the dephosphorylation of proteins modified by
XX phosphorylation of serine, threonine or tyrosine residues. The
XX phosphatases are useful for treating a variety of diseases: for example
XX cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,
XX synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
XX hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
XX glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
XX hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
XX congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
XX ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
XX Zonana syndrome, schizophrenia and hamartomas.
XX
XX Sequence 150 AA:

```

```

Query Match 96.4%; Score 769; DB 22; Length 150;
Best Local Similarity 95.3%; Pred. No. 1e-81;
Matches 143; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 MGVPNPNFNSWVLPGRLAGLALPRLPAAHYQFLDGVHIVSLTERGPHSDSCGTLHR 60

```

```
DB 1 MGVOPEFNSWVLPGRGLALPRLPAHYQFLDGVRLVSLTERGPHSDSCGLTLHR 60
QY 61 LRIIDFPCPPADQIDRFVQIVDEANANGAVGHCALGFGRGTMLACYLKERGLAAGD 120
DB 61 MRIPDFCPSPEQIDQFVKIYDEANANGAVGHCALGFGRGTMLACYLKERGLAAGD 120
QY 121 AIAEIRRLRPGSITETYEQEKAVFOFYQRTK 150
DB 121 AIAEIRRLRPGSITETYEQEKAVFOFYQRTK 150

RESULT 9
AAB35276
ID AAB35276 standard; Protein; 150 AA.
AC AAB35276;
XX
DT 08-MAY-2001 (first entry)
DE Murine dual specificity phosphatase DSP-11.
XX
KW Mouse; dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
KM MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
KM graft-versus-host disease; autoimmune disease; metabolic disease.
XX
OS Mus sp.
XX
PN WO200105983-A1.
XX
PD 25-JAN-2001.
XX
PF 19-JUL-2000; 2000WO-0519710.
XX
PR 20-JUL-1999; 99US-0144557.
XX
PA (CEPT-) CEPHYR INC.
XX
PI Lucbe RM, Wei B;
XX
DR WPI; 2001-147348/15.
XX
DR N-PSDB; AAF27959.
XX
PT Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase, is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
PS Example 3; Fig 6; 65pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human dual specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signalling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC mouse DSP-11 protein.
XX
SQ Sequence 150 AA;

Query Match 95.7%; Score 764; DB 22; Length 150;
Best Local Similarity 94.7%; Pred. No. 3.9e-81;
Matches 142; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVOPEFNSWVLPGRGLALPRLPAHYQFLDGVRLVSLTERGPHSDSCGLTLHR 60
DB 1 MGVOPEFNSWVLPGRGLALPRLPAHYQFLDGVRLVSLTERGPHSDSCGLTLHR 60
QY 61 LRIIDFPCPPADQIDRFVQIVDEANANGAVGHCALGFGRGTMLACYLKERGLAAGD 120
DB 61 MRIPDFCPSPEQIDQFVKIYDEANANGAVGHCALGFGRGTMLACYLKERGLAAGD 120
QY 121 AIAEIRRLRPGSITETYEQEKAVFOFYQRTK 150
DB 121 AIAEIRRLRPGSITETYEQEKAVFOFYQRTK 150
```

```
DB 121 AIAEIRRLRPGSITETYEQEKAVFOFYQRTK 150
QY 121 AIAEIRRLRPGSITETYEQEKAVFOFYQRTK 150

RESULT 10
ABB07847
ID ABB07847 standard; Protein; 681 AA.
AC ABB07847;
XX
DT 03-JUL-2002 (first entry)
DE C. elegans protein identified by EMBL Accession No. AF000363.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antirhythmic; hypotensive; vulnerery.
XX
OS Caenorhabditis elegans.
XX
PN WO200220747-A2.
XX
PD 14-MAR-2002.
XX
PF 05-SEP-2001; 2001WO-EP10205.
XX
PR 11-SEP-2000; 2000US-231568P.
PR 06-DEC-2000; 2000US-251403P.
XX
PA (FARB ) BAYER AG.
XX
PI Kossida S;
XX
DR WPI; 2002-339803/37.
XX
PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases -
XX
PS Disclosure; Fig 3; 117pp; English.
XX
CC The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischaemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. The present sequence
CC represents a C. elegans protein identified by EMBL Accn No. AF000363,
CC used for alignment studies with the human tyrosine phosphatase-like
CC enzyme polypeptide.
XX
SQ Sequence 681 AA;

Query Match 23.4%; Score 186.5; DB 23; Length 681;
Best Local Similarity 29.8%; Pred. No. 9.1e-13;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

QY 3 VQPPNFSWVLPGRGLALPRLPAHYQFLDGVRLVSLTERGPHSDSCGLTLHR 42
DB 192 VENGDFNWIIPGKITISFCGPHNESRENGPYHAPDYVDFYREKXKSTIVLAKKNYA 251
```

OY 43 ---TERGPHSDSCPGTLRLRIPDCPPAPDIDRFVQIVDEANMGEAVGVHCAIGF 99
 Db 252 SKFTKAGFDHVD-----LFTIDGSTPDEIMLKFIKIVD--NRKG-GVAHVCKRAGL 299
 OY 100 GRGTGTLACTIVKRGGLAGDAIEIRLRPGST----ETYEOKAVF 143
 Db 300 GRGTGTLIACMMKMEYGLTAGECMGLVRCRGSYIGPQPYLIEKOF 347

RESULT 11

AA67453
 ID AAG67453 standard; Protein: 447 AA.

AC AAG67453;

DT 26-NOV-2001 (first entry)

DE Amino acid sequence of a human polypeptide.

KW Human; protein kinase; protein phosphatase; signal transduction;

OS Homo sapiens.

PN WO200109345-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000MO-JP05060.

PR 29-JUL-1999; 99JP-0248036.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;

PI Senoo C, Nezu J;

DR WPI; 2001-564736/63.

XX New genes encoding protein kinase and protein phosphatase, useful for

XX identifying modulators which can be used to treat human or animal

XX disorders associated with the expression or function of these enzymes -

XX Example 4; Page 291-294; 336pp; Japanese.

XX The specification describes human protein kinase/protein phosphatases.

XX CC in cells. The kinase phosphatases are connected with intracellular

XX CC signaling pathways. Antisense oligonucleotides and compounds

XX CC identified by screening (agonists or antagonists) can be used to

XX CC treat human or animal disorders associated with the expression

XX CC or function of the protein. In addition, the polypeptides may be used

XX CC as target molecules for drug development. The present sequence

XX CC represents a polypeptide, used in the course of the invention.

XX SQ Sequence 447 AA;

Query Match 21.4%; Score 170.5; DB 22; Length 447;

Best Local Similarity 29.4%; Pred. No. 3.9e-11; Indels 31; Gaps 7;

Matches 50; Conservative 27; Mismatches 62;

OY 7 NFSWVLPGRGLALP---RLPAHY-----QFLDVGRLVSLTER---GPPHS 50

Db 185 DLWMIIPDRIRACGPHSRARLESQYHSHPEYTIQYFKHNHTTTIRLKKRYDAKRF 244

OY 51 DSCPGTLRLRLRIPDCPPAPDIDRFVQIVDEANMGEAVGVHCAIGFRTGTMACYL 110

Db 245 DA--GFDDHDLFPADGSTPTDAIVKRFIDICE--NAEG-ALAVHCKRAGLRTGTLACYI 299

Db 245 DA--GFDDHDLFPADGSTPTDAIVKRFIDICE--NAEG-ALAVHCKRAGLRTGTLACYI 299
 OY 111 VKRGLAAGDAIAELRLRPGSIETYEOKAVFO-----FYQRTK 150
 Db 300 MKHYRMTAETIIVWYRICRPLVIGPOQFLVMKOTSLMBGDFRQRLK 349

RESULT 12

AA67632
 ID AAG67632 standard; Protein: 447 AA.

AC AAG67632;

DT 26-NOV-2001 (first entry)

DE Amino acid sequence of a human protein.

KW Human; protein kinase; protein phosphatase; signal transduction.

OS Homo sapiens.

PN WO200109316-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000MO-JP05061.

PR 29-JUL-1999; 99JP-0248036.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;

PI Senoo C, Nezu J;

DR WPI; 2001-570286/64.

XX New genes encoding proteins with protein kinase/protein phosphatase

XX activity, useful in the diagnosis and treatment of diseases -

XX Example 4; Page 191-194; 233pp; Japanese.

XX The specification describes human protein kinase/protein phosphatases.

XX CC It is expected that the protein kinase/protein phosphatase gene

XX CC participates in signal transduction in cells. The protein

XX CC kinase/protein phosphatase polypeptides and polynucleotides are

XX CC useful for developing diagnostics and treatment agents for human

XX CC and animal diseases. The protein kinase/protein phosphatase polypeptides

XX CC are useful as target molecules in designing novel drugs. The protein

XX CC kinase/protein phosphatase polynucleotides are useful as a source of

XX CC probes and primers, which may be used to isolate homologous sequences.

XX CC The present sequence represents a human protein, which is used in the

XX SQ Sequence 447 AA;

Query Match 21.4%; Score 170.5; DB 22; Length 447;

Best Local Similarity 29.4%; Pred. No. 3.9e-11; Indels 31; Gaps 7;

Matches 50; Conservative 27; Mismatches 62;

OY 7 NFSWVLPGRGLALP---RLPAHY-----QFLDVGRLVSLTER---GPPHS 50

Db 185 DLWMIIPDRIRACGPHSRARLESQYHSHPEYTIQYFKHNHTTTIRLKKRYDAKRF 244

OY 51 DSCPGTLRLRLRIPDCPPAPDIDRFVQIVDEANMGEAVGVHCAIGFRTGTMACYL 110

Db 245 DA--GFDDHDLFPADGSTPTDAIVKRFIDICE--NAEG-ALAVHCKRAGLRTGTLACYI 299

OY 111 VKERGLAAGDAIAEIRRLPGSIENYEOEKAVFO-----FYORFK 150
DB 300 MKHRTMRAETITAWVRICRPGVLGPGQOFTLVKOTSLMEGDIYRORLK 349

RESULT 13
AAU75362
ID AAU75362 standard; Protein: 578 AA.

AC AAU75362;
DT 09-APR-2002 (first entry)
XX

DE Human dual specificity phosphatase CDC14A deletion variant.
XX
XX Human; cell-cycle control; CDC14A; cancer; deletion variant;
KW prostate cancer; breast cancer; tumour; lymph node metastasis;
KW malignant mesothelioma; chromosome 1p21; dual specificity phosphatase;
KW gene therapy; protein replacement therapy.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Misc-difference 417..418
XX /note="Wild-type RSDPTKGRHRAVSQPPRL substituted by RL"

US6331614-B1.
PD 18-DEC-2001.
PE 22-DEC-1999; 99US-0468872.
PR 23-DEC-1998; 98US-113833P.
XX (MYRI-) MYRIAD GENETICS INC.
XX
XX Wong AKC, Teng DHF, Tavtigian SV;
PI WPI; 2002-129551/17.
DR N-PSDB; ABK13395.
XX

PT Nucleic acid encoding mutated form of human dual-specificity
PT phosphatase CDC14A polypeptide, useful to diagnose and treat cancers
XX
XX Example 1; Page -: 41pp; English.

PS The invention relates to an isolated nucleic acid encoding a CDC14A
XX polypeptide (cell-cycle control protein 14A, a dual specificity
XX phosphatase), its complement or RNA molecule corresponding to it.
XX Also included are an expression vector comprising the nucleic acid
XX and a host cell transformed with the vector. The gene for CDC14A is
XX located on human chromosome 1p21. The nucleic acid and protein are useful
XX to diagnose and treat human cancers (e.g. breast cancer, prostate
XX cancer) and tumours (e.g. lymph node metastasis, malignant mesothelioma)
XX which have a mutation in the CDC14A gene, by gene therapy, protein
XX replacement therapy or protein mimetics. They can also be used to
XX screen for drugs to treat cancer. The present sequence represents a
XX deletion variant of CDC14A encoded by a cDNA isolated from a breast cell
XX line.
XX Note: The present sequence is not shown in the specification but was
XX created by the indexer using the CDC14A sequence appearing as AAU75361
XX and the information in table 4.

SQ Sequence 578 AA;

Query Match 20.1%; Score 160.5; DB 23; Length 578;
Best Local Similarity 29.9%; Pred. No. 8.1e-10;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

OY 3 VQPPNFSWVLPGRLA-----GLALPRLPAHYQFLDLGVNHLVSL----- 42
DB 175 VENGDFMWTIVGKRLASGPHPKSKIENGYPDLAEPAYFFYFKKHNTAVVRLNKKIYEA 234

OY 43 ---TERGPPHSDSCGLTLHRLRIPDFCPAPDQIDRFVQIVDEANANGAVGHCALGF 99
DB 235 KRFTDAGEHND-----LFFIDGSTPSPDNIVRRLNICE--NTEG-AIAVHCKAGL 282

OY 100 GRTGTMLACYLVKERGLAAGDAIAEIRRLRPGSI 133
DB 283 GRTGTMLACYLVKRTFTHAETIAWIRICRPGSI 316

RESULT 14
AAU75361
ID AAU75361 standard; Protein: 594 AA.

AC AAU75361;
DT 09-APR-2002 (first entry)
XX

DE Human dual specificity phosphatase CDC14A.
XX
XX Human; cell-cycle control; CDC14A; cancer;
KW prostate cancer; breast cancer; tumour; lymph node metastasis;
KW malignant mesothelioma; chromosome 1p21; dual specificity phosphatase;
KW gene therapy; protein replacement therapy.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Misc-difference 417..418
XX /note="Wild-type RSDPTKGRHRAVSQPPRL substituted by RL"

US6331614-B1.
PD 18-DEC-2001.
PE 22-DEC-1999; 99US-0468872.
PR 23-DEC-1998; 98US-113833P.
XX (MYRI-) MYRIAD GENETICS INC.
XX
XX Wong AKC, Teng DHF, Tavtigian SV;
PI WPI; 2002-129551/17.
DR N-PSDB; ABK13392.
XX

PT Nucleic acid encoding mutated form of human dual-specificity
PT phosphatase CDC14A polypeptide, useful to diagnose and treat cancers
XX
XX Claim 1; Column 58-62; 41pp; English.

PS The invention relates to an isolated nucleic acid encoding a CDC14A
XX polypeptide (cell-cycle control protein 14A, a dual specificity
XX phosphatase), its complement or RNA molecule corresponding to it.
XX Also included are an expression vector comprising the nucleic acid
XX and a host cell transformed with the vector. The gene for CDC14A is
XX located on human chromosome 1p21. The nucleic acid and protein are useful
XX to diagnose and treat human cancers (e.g. breast cancer, prostate
XX cancer) and tumours (e.g. lymph node metastasis, malignant mesothelioma)
XX which have a mutation in the CDC14A gene, by gene therapy, protein
XX replacement therapy or protein mimetics. They can also be used to
XX screen for drugs to treat cancer. The present sequence represents CDC14A.
XX

SQ Sequence 594 AA;

Query Match 20.1%; Score 160.5; DB 23; Length 594;
Best Local Similarity 29.9%; Pred. No. 8.4e-10;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

OY 3 VQPPNFSWVLPGRLA-----GLALPRLPAHYQFLDLGVNHLVSL----- 42
DB 175 VENGDFMWTIVGKRLASGPHPKSKIENGYPDLAEPAYFFYFKKHNTAVVRLNKKIYEA 234

OY 43 ---TERGPPHSDSCGLTLHRLRIPDFCPAPDQIDRFVQIVDEANANGAVGHCALGF 99
DB 235 KRFTDAGEHND-----LFFIDGSTPSPDNIVRRLNICE--NTEG-AIAVHCKAGL 282

OY 100 GRTGTMLACYLVKERGLAAGDAIAEIRRLRPGSI 133

Db 283 GRGTGLIACYMKHYRFTHAETIAMIRICRPGSI 316

RESULT 15

AAU75363

ID AAU75363 standard; Protein: 594 AA.

AC AAU75363;

DT 09-APR-2002 (first entry)

DE Human dual specificity phosphatase CDC14A G571R variant.

XX Human; cell-cycle control; CDC14A; cancer; G571R;

KW prostate cancer; breast cancer; tumour; lymph node metastasis;

KM malignant mesothelioma; chromosome 1p21; dual specificity phosphatase;

XX gene therapy; protein replacement therapy.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 571

PD 18-DEC-2001.

PF 22-DEC-1999; 99DS-0468872.

PR 23-DEC-1998; 98DS-113833P.

PA (MYRI-) MYRIAD GENETICS INC.

PI Wong AKC, Teng DHF, Tavtigian SV;

DR WPI: 2002-129551/17.

DR N-PSDB; ABK13397.

XX Nucleic acid encoding mutated form of human dual-specificity phosphatase CDC14A polypeptide, useful to diagnose and treat cancers -

XX Example 1; Page -: 41pp; English.

CC The invention relates to an isolated nucleic acid encoding a CDC14A polypeptide (cell-cycle control protein 14A, a dual specificity

CC phosphatase), its complement or RNA molecule corresponding to it.

CC Also included are an expression vector comprising the nucleic acid

CC and a host cell transformed with the vector. The gene for CDC14A is

CC located on human chromosome 1p21. The nucleic acid and protein are useful

CC to diagnose and treat human cancers (e.g. breast cancer, prostate

CC cancer) and tumours (e.g. lymph node metastasis, malignant mesothelioma)

CC which have a mutation in the CDC14A gene, by gene therapy, protein

CC replacement therapy or protein mimetics. They can also be used to

CC screen for drugs to treat cancer. The present sequence represents a

CC variant of CDC14A.

CC Note: The present sequence is not shown in the specification but was

CC created by the indexer using the CDC14A sequence appearing as AAU75361

CC and the information in table 4.

XX SQ Sequence 594 AA;

XX Query Match 20.1%; Score 160.5; DB 23; Length 594;

XX Best Local Similarity 29.9%; Pred. No. 8.4e-10;

XX Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

QY 3 VDPNPSWVLPGRIA-----GLAPRLPAHYQFLDLGVRHLVSL----- 42

DB 175 VENGDPMNIVPGKFLAFSGHPKSKIENGYPHAPAEAFYFKKNTATAYVRLNKITYEA 234

QY 43 ----TENGPHSDSCPGILTLRLRIIPDFCPAPQIDRFVQIVDEANRGEAVGVHCALGF 99

Db 235 KRFTDAGEFHYD-----LFFIDGSTPSDNIVRRLINICE--NTEG-AIAYHCKAGL 282

QY 100 GRGTGLIACYMKHYRFTHAETIAMIRICRPGSI 133

DB 283 GRGTGLIACYMKHYRFTHAETIAMIRICRPGSI 316

Search completed: November 4, 2002, 12:46:16
Job time : 69 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 4, 2002, 12:44:52 ; Search time 26 seconds
(without alignments)
169.748 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798
Sequence: 1 MGVPNPNFSWVLPGRILA.....GSEIYEOEKAVFOFYORTK 150

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160.5	20.1	594	US-09-468-872-2	Sequence 2, Appl
2	128.5	16.1	196	US-09-752-165-95	Sequence 95, Appl
3	117.5	14.7	173	US-09-704-139-4	Sequence 4, Appl
4	115	14.4	195	US-09-752-165-93	Sequence 93, Appl
5	115	14.4	195	US-09-752-165-96	Sequence 96, Appl
6	115	14.4	597	US-09-134-218-4	Sequence 4, Appl
7	115	14.4	597	US-09-134-218-5	Sequence 5, Appl
8	114	14.3	168	US-09-188-579-85	Sequence 85, Appl
9	114	14.3	168	US-09-315-444-85	Sequence 85, Appl
10	114	14.3	168	US-09-721-362-85	Sequence 85, Appl
11	114	14.3	168	US-09-752-165-99	Sequence 99, Appl
12	113	14.2	197	US-09-752-165-94	Sequence 94, Appl
13	112	14.0	173	US-09-188-579-87	Sequence 87, Appl
14	112	14.0	173	US-09-315-444-87	Sequence 87, Appl
15	112	14.0	173	US-09-721-362-87	Sequence 87, Appl
16	111.5	14.0	353	US-09-013-881-3	Sequence 3, Appl
17	111	13.9	172	US-09-704-139-5	Sequence 5, Appl
18	110	13.8	201	US-08-933-750C-21	Sequence 21, Appl
19	110	13.8	201	US-09-234-613-21	Sequence 21, Appl
20	107	13.4	176	US-09-704-139-2	Sequence 2, Appl
21	105.5	13.2	453	US-08-951-260A-7	Sequence 7, Appl
22	104	13.0	263	US-08-685-992-13	Sequence 13, Appl
23	104	13.0	263	US-09-144-925-13	Sequence 13, Appl
24	104	13.0	595	US-08-202-389-6	Sequence 6, Appl
25	104	13.0	631	US-08-202-389-8	Sequence 8, Appl
26	103.5	13.0	453	US-08-821-278A-2	Sequence 2, Appl
27	102.5	12.8	173	US-09-164-193-2	Sequence 2, Appl

28	102.5	12.8	173	US-09-221-448A-2	Sequence 2, Appl
29	102	12.8	175	US-09-188-579-86	Sequence 86, Appl
30	102	12.8	175	US-09-315-444-86	Sequence 86, Appl
31	102	12.8	175	US-09-721-362-86	Sequence 86, Appl
32	102	12.8	513	US-08-202-389-2	Sequence 2, Appl
33	101.5	12.7	369	US-08-596-291-4	Sequence 4, Appl
34	101.5	12.7	369	US-09-100-804-4	Sequence 4, Appl
35	101.5	12.7	369	US-09-100-804-4	Sequence 4, Appl
36	98.5	12.3	699	US-08-148-006B-7	Sequence 7, Appl
37	98.5	12.3	699	US-08-800-825A-7	Sequence 7, Appl
38	98.5	12.3	699	US-09-158-657-7	Sequence 7, Appl
39	98	12.3	249	US-08-685-992-28	Sequence 28, Appl
40	98	12.3	249	US-09-144-925-28	Sequence 28, Appl
41	98	12.3	360	US-09-020-743-2	Sequence 2, Appl
42	96.5	12.1	150	US-09-164-193-23	Sequence 23, Appl
43	96.5	12.1	150	US-09-221-448A-23	Sequence 23, Appl
44	95.5	12.0	528	US-08-987-691A-4	Sequence 4, Appl
45	94	11.8	278	US-08-821-278A-18	Sequence 18, Appl

ALIGNMENTS

```

RESULT 1
US-09-468-872-2
; Sequence 2, Application US/09468872
; Patent No. 6331614
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Teng, David H.F.
; APPLICANT: Tavligian, Sean V.
; TITLE OF INVENTION: Human CDC14A Gene
; FILE REFERENCE: CDC14A Gene
; CURRENT APPLICATION NUMBER: US/09/468,872
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: US 60/113,833
; EARLIER FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 2
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-468-872-2

Query Match      20.1%; Score 160.5; DB 4; Length 594;
Best Local Similarity 29.9%; Pred. No. 2.7e-11;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

QY      3  VQPNPNSWVLPGRILA-----GLAUPRLRAHYQFLDLGVRHLYL----- 42
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 175  VENGDDNVIYVPGKFLAFSGPHPSKIENGLPLHAPRPYFFYFKHVTAVRINKKIYEA 234
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 43  ---TERGPHSDSCPGILTLRLRIPDCPPAPQIDRFVOYDEANARGEAVGHCALGF 99
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 235  KRRTDAGFEHYD-----LEFIDGSPSDIVIRFLNICE--NTEG-AINHVCKAGL 282
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY      100 GRGTGLACIYVKEKRLAAGDAIAETRLRPGSI 133
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 283  GRGTGLACIYVKNRYFTFAETIRICRPGSI 316
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 2
US-09-752-165-95
; Sequence 95, Application US/09752165
; Patent No. 6451583
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; APPLICANT: Ho, C. Kiong
; TITLE OF INVENTION: Pharmacological Targeting Of mRNA Cap Formation
; TITLE OF INVENTION: For treatment of Parasitic Infections
; FILE REFERENCE: D6388
; CURRENT APPLICATION NUMBER: US/09/752,165

```

RESULT 6
 US-09-134-218-4
 Sequence 4, Application US/09134218B
 Patent No. 6532926
 GENERAL INFORMATION:
 APPLICANT: Shattin, Aaron J.
 APPLICANT: Applicant, Renka
 APPLICANT: Reinders, Danny
 APPLICANT: Yu, Zheng
 APPLICANT: Moldanado, Edlio

```

; TITLE OF INVENTION: mRNA CAPPING ENZYMES AND USES THEREOF
; FILE REFERENCE: 601-1-079 ss
; CURRENT APPLICATION NUMBER: US/09/134,218A
; CURRENT FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-134-218-4

Query Match
Best local Similarity 30.0%; Pred. No. 1.2e-05;
Matches 24; Conservative 18; Mismatches 36; Indels 2; Gaps 1;

QY 69 PARDQIDRFVQIVDEANARG--EAVGVHCAAGFGRTGTMACYIKERGLAAGDAIAEIR 126
Db 98 PTENTETFTRLCEERFENRPPPELLIGVHCTHGFNRTGFLICAPLVERKMDWSIEAAVATFA 157

QY 127 RLRRGSIETYEOKAVFOFY 146
Db 158 QARPPGIYKGDYKELFRRY 177

RESULT 7
US-09-134-218-5
; Sequence 5, Application US/09134218A
; Patent No. 6312926
; GENERAL INFORMATION:
; APPLICANT: Shatkin, Aaron J.
; APPLICANT: Pillutla, Renuka
; APPLICANT: Reinberg, Danny
; APPLICANT: Yu, Zheng
; APPLICANT: Moldonado, Edio
; TITLE OF INVENTION: mRNA CAPPING ENZYMES AND USES THEREOF
; FILE REFERENCE: 601-1-079 ss
; CURRENT APPLICATION NUMBER: US/09/134,218A
; CURRENT FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-134-218-5

Query Match
Best local Similarity 30.0%; Pred. No. 1.2e-05;
Matches 24; Conservative 18; Mismatches 36; Indels 2; Gaps 1;

QY 69 PARDQIDRFVQIVDEANARG--EAVGVHCAAGFGRTGTMACYIKERGLAAGDAIAEIR 126
Db 98 PTENTETFTRLCEERFENRPPPELLIGVHCTHGFNRTGFLICAPLVERKMDWSIEAAVATFA 157

QY 127 RLRRGSIETYEOKAVFOFY 146
Db 158 QARPPGIYKGDYKELFRRY 177

RESULT 8
US-09-188-579-85
; Sequence 85, Application US/09188579B
; Patent No. 6107040
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6185
; CURRENT APPLICATION NUMBER: US/09/188,579B
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO 85
; LENGTH: 168
```

```

; TYPE: PRT
; ORGANISM: baculovirus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of RNA baculovirus phosphatase BVP.
US-09-188-579-85

Query Match
Best local Similarity 32.6%; Pred. No. 2.7e-06;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;

QY 55 GLTLHRLRIDFPCPPARDQIDRFVQIVDEANAR--GEAVGVHCAAGFGRTGTMACYLYK 112
Db 77 GLYKKIQQVGQTLPPESIVQEFIDYVKEFTKCPGMLVGVHCTHGINRGYMCRYLMH 136

QY 113 ERGLAAGDAIAEIRRLRRPSIE--TYEOE 139
Db 137 TLGIAPOEALIDREFEKAHGKHIERONYOD 165

RESULT 9
US-09-315-444-85
; Sequence 85, Application US/09315444A
; Patent No. 6232070
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6185CIP
; CURRENT APPLICATION NUMBER: US/09/315,444A
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 09/188,579
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 85
; LENGTH: 168
; TYPE: PRT
; ORGANISM: baculovirus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of RNA baculovirus phosphatase BVP.
US-09-315-444-85

Query Match
Best local Similarity 32.6%; Pred. No. 2.7e-06;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;

QY 55 GLTLHRLRIDFPCPPARDQIDRFVQIVDEANAR--GEAVGVHCAAGFGRTGTMACYLYK 112
Db 77 GLYKKIQQVGQTLPPESIVQEFIDYVKEFTKCPGMLVGVHCTHGINRGYMCRYLMH 136

QY 113 ERGLAAGDAIAEIRRLRRPSIE--TYEOE 139
Db 137 TLGIAPOEALIDREFEKAHGKHIERONYOD 165

RESULT 10
US-09-721-362-85
; Sequence 85, Application US/09721362
; Patent No. 6420163
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6185CIP/D
; CURRENT APPLICATION NUMBER: US/09/721,362
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 09/315,444
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 85
; LENGTH: 168
; TYPE: PRT
; ORGANISM: baculovirus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of RNA baculovirus phosphatase BVP.
US-09-721-362-85
```

```

Query Match Similarity      14.3%; Score 114; DB 4; Length 168;
Best Local Similarity      32.6%; Pred. No. 2.7e-06;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;

QY 55 GUTLRLRIPDFCPAPDQIDRFVQIYDEANAR--GEAVGYHCALGRTGTMLACTYK 112
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 77 GLTYKKIIVPGQTLRPESIVGEFIDTYKEFTKCPGMLVGHCHGINRTGYWCRYLHM 136
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 113 ERGLAAGDAIAEIRLRPGSIE--TYEQE 139
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 137 TLGIAPGEAIDRFKARGHKIERQNYQD 165
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 11
US-09-752-165-99
; Sequence 99, Application US/09752165
; Patent No. 6451583
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; APPLICANT: Ho, C. Kiong
; TITLE OF INVENTION: Pharmacological Targeting Of mRNA Cap Formation
; TITLE OF INVENTION: For Treatment Of Parasitic Infections
; FILE REFERENCE: D6388
; CURRENT APPLICATION NUMBER: US/09/752,165
; CURRENT FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 104
; SEQ ID NO 99
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Baculovirus RNA-specific 5' phosphatase BVP
US-09-752-165-99

Query Match      14.3%; Score 114; DB 4; Length 168;
Best Local Similarity 32.6%; Pred. No. 2.7e-06;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;

QY 55 GUTLRLRIPDFCPAPDQIDRFVQIYDEANAR--GEAVGYHCALGRTGTMLACTYK 112
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 77 GLTYKKIIVPGQTLRPESIVGEFIDTYKEFTKCPGMLVGHCHGINRTGYWCRYLHM 136
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 113 ERGLAAGDAIAEIRLRPGSIE--TYEQE 139
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 137 TLGIAPGEAIDRFKARGHKIERQNYQD 165
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 12
US-09-752-165-94
; Sequence 94, Application US/09752165
; Patent No. 6451583
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; APPLICANT: Ho, C. Kiong
; TITLE OF INVENTION: Pharmacological Targeting Of mRNA Cap Formation
; TITLE OF INVENTION: For Treatment Of Parasitic Infections
; FILE REFERENCE: D6388
; CURRENT APPLICATION NUMBER: US/09/752,165
; CURRENT FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 104
; SEQ ID NO 94
; LENGTH: 197
; TYPE: PRT
; ORGANISM: C elegans
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: RNA triphosphatase domain of C elegans RNA capping enzyme
US-09-752-165-94

Query Match      14.2%; Score 113; DB 4; Length 197;
Best Local Similarity 36.0%; Pred. No. 4.4e-06;

```

```

Matches      31;   Conservative    12;   Mismatches     37;   Indels        6;   Gaps         3

Oy       69  PAPDIDDFVQIVYDDANR--GEANGVCALGFGRTMLACYL--VKERGLAAGDAIAE 124
           | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db        95  PLEQTDNFIKLVQEFHKKRPDRVVGVGHCTHGTFNLAAVLQVEVEYGDA--AIGE 152
           | : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy       125  IRRLRPGSIETPEGEKAVGFQRFQRK 150
           | : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db        153  FAENRKGIYKODYTDDLFAKYDPTE 178

RESULT 13
US-09-188-579-87
; Sequence 87, Application US/09188579B
; Patent No. 6107040
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6185
; CURRENT APPLICATION NUMBER: US/09/188,579B
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO 87
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of N-terminal RNA triphosphatase
US-09-188-579-87
```

```

OY      69  PAPDQIDRFVQIVDEANR--GEAVGVHCALGFGRGTMLACYL--VKERGLAAGDAIAE 124
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      96  PTOEDTDNFILKVOEFHKKYDPDRVGVGCHTGFNRTGFLIAAYLQVEVEYGIDA--AIGE 153
Matches 30; Conservative 11; Mismatches 26; Indels 8; Gaps 4;

OY      125  IRLRPGSIETYEQE 139
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      154  FAENRQKGI--YKOD 166

RESULT 14
US-09-315-444-87
: Sequence 87, Application US/09315444A
: Patent No. 6232070
: GENERAL INFORMATION:
: APPLICANT: Shuman, Stewart
: TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
: FILE REFERENCE: D6185CIP
: CURRENT APPLICATION NUMBER: US/09/315,444A
: CURRENT FILING DATE: 1999-05-20
: PRIOR APPLICATION NUMBER: US 09/188,579
: PRIOR FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 87
: LENGTH: 173
: TYPE: PRT
: ORGANISM: Candida albicans
: FEATURE:
: OTHER INFORMATION: Amino acid sequence of N-terminal RNA triphosphatase
US-09-315-444-87

OY      69  PAPDQIDRFVQIVDEANR--GEAVGVHCALGFGRGTMLACYL--VKERGLAAGDAIAE 124
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      96  PTOEDTDNFILKVOEFHKKYDPDRVGVGCHTGFNRTGFLIAAYLQVEVEYGIDA--AIGE 153
Matches 30; Conservative 11; Mismatches 26; Indels 8; Gaps 4;

```

```

OY      125  IRLRPGSIETYEQE 139
          |      | : | :
Db      154  FAENRQKGI - YKQD 166

```

RESULT 15
HE-00-731

US-09-721-362-87

Sequence 87, Application US/09721362

Patent NO. 6420163

; GENERAL INFORMATION:

APPLICANT: Shuman, Stewart

TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
FILE REFERENCE: D6185CTD/D

FILE REFERENCE: D6185C1P/D

; CURRENT APPLICATION NUMBER: US/09/121,362
 ; CURRENT FILING DATE: 2000-11-23

;; CURRENT FILING DATE: 2000-11-22
;; PRIOR APPLICATION NUMBER: US 09

PRIOR FILING DATE: 1999-05-20

NUMBER OF SEQ ID NOS: 116

; NUMBER OF SEQ ID NOS. 110
; SEQ ID NO 87 .

; LENGTH: 17

TYPE: PRT

ORGANISM

FEATURE:

OTHER IN

OTHER INFORMATION: domain of the capping enzyme.

US-09-721-362-87

1

Query Match

Best Local Similarity 40.08; Pred. No. 4.9e-06;

Matches 30; Conservative 11; Mismatches 26; Indels 8; Gaps 4;

Qy	69	PAP0IDFV0IVDEANAR--GEAVGHCAAGFERTGTMACYL--VKESGLAGDIAIE	124
Db	96	PTQEDTDMFNLVQGEFHKKPYDRVVGVCSTFGFNRIGFLIAAYLQFYEBVGDLA--AIGE	133
Qy	125	IRRLRPGSIETYEDE	139
Db	154	FAENRQKGI--YKOD	166

Search completed: November 4, 2002, 12:50:38
Job time : 28 secs


```

: Sequence 4, Application US/09815419
: Patent No. US20020065406A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel A.
: TITLE OF INVENTION: 18221, A NOVEL DUAL SPECIFICITY
: TITLE OF INVENTION: PHOSPHATASE AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: 10448-030001
: CURRENT APPLICATION NUMBER: US/09/815,419
: CURRENT FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: US 60/191,858
: PRIOR FILING DATE: 2000-03-24
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 173
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus sequence
: US-09-815-419-4

```

[illegible]

```

: RESULT 3
: US-09-964-277-7
: Sequence 7, Application US/09964277
: Patent No. US20020137170A1
: GENERAL INFORMATION:
: APPLICANT: Lucne, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
: FILE REFERENCE: 200125.434
: CURRENT APPLICATION NUMBER: US/09/964,277
: CURRENT FILING DATE: 2001-09-25
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 155
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-964-277-7

```

Query Match	14.3%	Score 114	DB 10	Length 155
Best Local Similarity	26.4%	Pred. No. 2.2e-05		
Matches	28	Conservative	20	Mismatches 42; Indels 16; Gaps 3;
OY	50	SDSCPELT/L-----HRLRIP-----DECPAPQDIREVQIVDEANRGEAVGHCALGFGRT	102	
DB	46	SNTCPPDPDIPESHFLRVPNVDSFCETILPWLKDSVDFETKAKNSGCVLHCHLAGISRS	105	
OY	103	GTMALCYLWKRGLAGDAIAEIRLRP-----GSIETFOE	139	
DB	106	ATATATATIKRMDMSLDEAATFYKERRPTISPNFNLGQLDIDDKK	151	

RESULT 4
US-09-964-277-21

```

: Sequence 21 Application US/09964277
: Patent No. US20020137170A1
:
: GENERAL INFORMATION:
:
: APPLICANT: LucHe, Ralf M.
:
: TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
: FILE REFERENCE: 200125.434
: CURRENT APPLICATION NUMBER: US/09/964,277
: CURRENT FILING DATE: 2001-09-25
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 21
:
: LENGTH: 517
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: US-09-964-277-21

```

	Query Match	14.3%	Score 114;	DB 10;	Length 517;
	Best Local Similarity	26.4%;	Pred. No. 0.0001;		
	Matches	28;	Conservative	20;	Mismatches 42; Indels 16; Gaps 3;
QY	50	SDSCPGILN----	HRLRIP-----	DCCPAPADODIRGVJVDENANGAEVGYHCAIGFRT	102
		::	:: :	:::	::
Db	44	SNTCKRPDPIFESHRLKPVANDSFCEKILLPWLDKSVDFTFKKAANGCVLHCLAGISRS			103
QY	103	GTMACLVKVERGLAADAIAEIRLRP-----	-SIEIYEOE		139
		::	:::	:	:: :
Db	104	ATIAIAYIMKRMDMSLDEAIRFVKCKRPTISPENFLQLQDLDYERK			149

```

RESULT 5
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-816-494-2

```

Query Match	14.3%	Score 114	DB 10	Length 665
Best Local Similarity	26.4%	Pred. No. 0.00014		
Matches 28	Conservative 20	Mismatches 42	Indels 16	Gaps 3

QY	50	SDSCPGTLT---	RRLRP---	DFCPAPQDIDRRVQYVDENMARGAVGVCHALGFERT	102
		: :	:	:	:
DB	192	SNTPCKPDPFIESHFLRPVNDSCFEKTLPMWIDSKSVDFIEKAKASNGCVLWCHLAGISRS	251		
		: :	:	:	:
QY	103	GTMLACVYFKERGLAGADAIAEIRLRP-----	GSIFTEYGE	139	
		: :	:	:	:
DB	252	ATIAIAITMKRDMDSLDEAYRPFVKERKRTISPENFETGLQLLDYEEK	297		
		: :	:	:	:

RESULT 6
 US-09-964-277-2
 : Sequence 2, Application US/09964277
 : Patent No. US20020137110A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: LucDe, Ralf M.
 :
 : APPLICANT: Mel, Bo
 :
 : FILE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
 : FILE REFERENCE: 200125.434

RESULT 15

Job time : 21 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 2002, 07:42:22 ; Search time 2785 Seconds

(without alignments)
7388.037 Million cell updates/sec

Title: US-09-619-380-1

Perfect score: 707
Sequence: 1 tgaccgcgtctccttgccc.....aaaaaaaaaaaaaaaaaaaaa 707Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
GenBank:
1: gb_ba:*
2: gb_mt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdt:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	707	100.0	707	6 AX078400	AX078400 Sequence
2	703.8	99.5	878	6 AX482482	AX482482 Sequence
3	703	99.4	839	6 AX099936	AX099936 Sequence
4	679.8	96.2	711	9 BC001140	BC001140 Homo sapi
5	673.6	95.3	746	6 AX398840	AX398840 Sequence
6	673.6	95.3	746	6 AX398843	AX398843 Sequence
7	673.6	95.3	746	6 AX398847	AX398847 Sequence
8	673.6	95.3	746	6 AX482336	AX482336 Sequence
9	671.6	95.0	709	6 AX398833	AX398833 Sequence
10	657.4	93.0	694	6 AX086042	AX086042 Sequence
11	583.8	82.6	602	6 AX398836	AX398836 Sequence
12	555.8	78.6	559	9 HMM2E12005	AX086042 Sequence
13	508.6	71.9	553	6 AX398834	AX398834 Sequence
14	486.2	68.8	433	6 AX398835	AX398835 Sequence
15	446	63.1	448	6 AX398830	AX398830 Sequence
16	429	60.7	453	6 AX086004	AX086004 Sequence
17	396.4	56.1	715	6 AL580560	AX086004 Sequence
18	391	55.3	13862	9 AL580560	AX086004 Sequence
19	379.8	53.7	397	6 AX398837	AX398837 Sequence
20	373	52.8	433	6 AX078411	AX078411 Sequence
21	336.4	47.6	399	6 AX398830	AX398830 Sequence
22	241.8	34.2	247	6 AX334867	AX334867 Sequence
23	229.6	32.5	17686	2 AC131177	AC131177 Mus muscu
24	80	11.3	80	6 AX482385	AX482385 Sequence
25	79	11.2	79	6 AX482382	AX482382 Sequence
26	71.8	10.2	35694	3 C1C0S41	AL096811 Streptomy
27	57.4	8.1	36583	1 SC5H1	AL096811 Streptomy
28	56.8	8.0	148766	2 AC118995	AL096811 Streptomy
29	56.2	7.9	35033	1 SC130A	AL096811 Streptomy
30	55.2	7.8	1146	6 I92666	192666 Sequence 1
31	55.2	7.8	2544	1 SA017169	U17169 Streptomyce
32	55.2	7.8	2728	6 AF429315	192670 Sequence 5
33	54.6	7.7	125020	9 AF429315	AF429315 Homo sapi
34	54	7.6	134350	2 AC105734	AC105734 Oryza sat
35	54	7.6	159636	8 AC090485	AC090485 Genomic S
36	52.8	7.5	40337	1 SC5A7	AC090485 Genomic S
37	52.8	7.5	110000	2 AC116301-0	AL031107 Streptomy
38	52.6	7.4	83582	2 AC121741	AC116301 Homo sapi
39	52.6	7.4	104627	2 AC117882	AC121741 Homo sapi
40	52.2	7.4	54750	2 AC009072	AC117882 Rattus no
41	52.2	7.4	214791	9 AC009072	AC009072 Homo sapi
42	52	7.4	79419	2 AC128943	AC007610 Homo sapi
43	51.2	7.2	69165	2 AC023465	AC128943 Rattus no
44	51	7.2	10961	1 AE001939	AC023465 Homo sapi
45	50.8	7.2	20021	1 AE004730	AE001939 Dendroccc
					AE004730 Pseudomon

ALIGNMENTS

RESULT 1	AX078400	707 bp	DNA	1 linear	PAT 22-FRB-2001
LOCUS	AX078400				
DEFINITION	Sequence 1 from Patent WO0105983.				
ACCESSION	AX078400				
VERSION	AX078400.1	GI:13158069			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 707)				
JOURNAL	Lucas, R.M. and Wei, B.				
	Dsp-11 dual-specificity map kinase phosphatase				
	Patent: WO 0105983-A 1 25-JAN-2001;				

FEATURES Ceptyr, Inc. (US)
Location/Qualifiers
Source 1..707
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 146 a 224 c 202 g 135 t
ORIGIN

Query Match 100.0%; Score 707; DB 6; Length 707;
Best Local Similarity 100.0%; Pred. No. 5.6e-117;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGGACACCCCACTTCCTCG 60
DB 1 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGGACACCCCACTTCCTCG 60
QY 61 GTGCTCCGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTTT 120
DB 61 GTGCTCCGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTTT 120
QY 121 CTGTTGACCTGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 180
DB 121 CTGTTGACCTGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 180
QY 181 GACAGCTGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTT 240
DB 181 GACAGCTGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTT 240
QY 241 CCGGACAGATGACCGCTGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTAC 300
DB 241 CCGGACAGATGACCGCTGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTAC 300
QY 301 GTGGAGTGCACCTGCTGCTGGGCTGGGAGTGGGCTGCGGCGCTCCCGCC 360
DB 301 GTGGAGTGCACCTGCTGCTGGGCTGGGAGTGGGCTGCGGCGCTCCCGCC 360
QY 361 GTGAAGAGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 420
DB 361 GTGAAGAGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 420
QY 421 GGCCTCATCAGACCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 GGCCTCATCAGACCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 TAAGGGGCTTGTACCTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 TAAGGGGCTTGTACCTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GCCAGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 GCCAGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 CACTGAAGTACCCAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 CACTGAAGTACCCAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GTGAATTAATGAGTTTACGAAACCAAAAAAAAAAAAAAAAAAAAAA 707
DB 661 GTGAATTAATGAGTTTACGAAACCAAAAAAAAAAAAAAAAAAAAAA 707

RESULT 2
AX482482 878 bp DNA linear PAT 16-AUG-2002
LOCUS Sequence 151 from Patent WO02057460.
DEFINITION AX482482
ACCESSION AX482482
VERSION AX482482.1 GI:22317003
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Bassolino,D., Feder,J., Krystek,S.,
Mcatee,P., Muller,G., Stiemers,N., Jackson,D.G., and Ramathan,C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 151 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES Location/Qualifiers
Source 1..878
/organism="Homo sapiens"
/db_xref="taxon:9606"
89..541
CDS /note="unnamed protein product"
/codon_start=1
/protein_id="CAD4455.1"
/db_xref="GI:22317004"
/translation="MGVQPPNFWLPLRLALPLRLPAHYOFLDGVHIVSLTE
RGPVHSDSCPELTHRLRLRIPDCEPAPQIDRFQVDEANARENAVGHALGFRT
GTMALCTVYKRGIAAGDAIAEIRLRPDSIETYEQEKAVFYQRTK"
BASE COUNT 259 a 247 c 233 g 139 t
ORIGIN

Query Match 99.5%; Score 703.8; DB 6; Length 878;
Best Local Similarity 99.7%; Pred. No. 2.1e-116;
Matches 705; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGGACACCCCACTTCCTCG 60
DB 59 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGGACACCCCACTTCCTCG 118
QY 61 GTGCTCCGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTTT 120
DB 119 GTGCTCCGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTTT 178
QY 121 CTGTTGACCTGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 180
DB 179 CTGTTGACCTGGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 238
QY 181 GACAGCTGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTT 240
DB 239 GACAGCTGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAGTT 298
QY 241 CCGGACAGATGACCGCTGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTAC 300
DB 299 CCGGACAGATGACCGCTGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTAC 358
QY 301 GTGGAGTGCACCTGCTGCTGGGCTGGGAGTGGGCTGCGGCGCTCCCGCC 360
DB 359 GTGGAGTGCACCTGCTGCTGGGCTGGGAGTGGGCTGCGGCGCTCCCGCC 418
QY 361 GTGAAGAGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 420
DB 419 GTGAAGAGCGGCGGCTGGGAGTGGGCTGCGGCGCTCCCGCCACTACAG 478
QY 421 GGCCTCATCAGACCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 479 GGCCTCATCAGACCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538
QY 481 TAAGGGGCTTGTACCTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 539 TAAGGGGCTTGTACCTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 598
QY 541 GCCAGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 599 GCCAGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
QY 601 CACTGAAGTACCCAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 659 CACTGAAGTACCCAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
QY 661 GTGAATTAATGAGTTTACGAAACCAAAAAAAAAAAAAAAAAAAAAA 707
DB 719 GTGAATTAATGAGTTTACGAAACCAAAAAAAAAAAAAAAAAAAAAA 765

Db	734	GTTGATATAATGAGTTTACGAACCAAAAAAAAAAAAAAAAAAAAAA	776
Result 4			
LOCUS	BC001140	711 bp	mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, hypothetical protein FLJ20442, clone MGC:10442		
ACCESSION	BC001140		
VERSION	BC001140.1	GI:12654608	
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 711)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: DCTD/DTF CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Fyell, Erin Garland, Ran Guin, Letlicia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pavan Pandoh, Anna-Elisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyerdunyn, Marco Marra.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 8 Row: 0 Column: 23 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7020544.		
SOURCE	Location/Qualifiers		
	1..711		
	/organism="Homo sapiens"		
	/db_xref="LocusID:54935"		
	/db_xref="taxon:9606"		
	/clone="MGC:1042 IMAGE:3535406"		
	/tissue_type="Lung, small cell carcinoma"		
	/clone_1lb="NIH-MGC_7"		
	/lab_host="DH10B-R"		
	/note="Vector: pOTB7"		
	21..473		
	/codon_start=1		
	/product="hypothetical protein FLJ20442"		
	/protein_id="AA01140.1"		
	/db_xref="GI:12654609"		
	/translation="MGVQVPSFMSVLPGRLAGIALPRLPAHYOFLDLGVRLVSLTE RGPSPSCPLTLHRLRPDPCPPADPDIDRFYQIYDEANAGEAAGVCAIGCFRT GTMLACTYVERGGAAGDAIAETIRLRPSGSIETVEDEKATFYQYRK"		
BASE COUNT	159 a 217 c 208 g 127 t		
ORIGIN			
Query Match	96.2%; Score 679.8; DB 9; Length 711;		
Best Local Similarity	99.7%; Pred. No. 4.1e-112;		
Matches 681; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		

ORIGIN

Query Match 95.3%; Score 673.6; DB 6; Length 746;
 Best Local Similarity 99.4%; Pred. No. 5.3e-111;
 Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 28 GCGATGGGGCTGCAGACCCCACTTCTCTGGGTGCTTCGCGGGCCGGCTGGCGGAGCTG 87
 |||||||
 Db 32 GCGATGGGGCTGCAGACCCCACTTCTCTGGGTGCTTCGCGGGCCGGCTGGCGGAGCTG 91
 |||||||

QY 88 GCGGTGCGCGGCTGCCCGCCCACTACAGAGTTCCTGTTGGACCTGGGGCTGGCGGAGCTG 147
 |||||||
 Db 92 GCGGTGCGCGGCTGCCCGCCCACTACAGAGTTCCTGTTGGACCTGGGGCTGGCGGAGCTG 151
 |||||||

QY 148 GTGTCCCTGACGAGCGGGGCGCCCTTCACAGCGACGTGCTCCCGGCTTCACCTGAC 207
 |||||||
 Db 152 GTGTCCCTGACGAGCGGGGCGCCCTTCACAGCGACGTGCTCCCGGCTTCACCTGAC 211
 |||||||

QY 208 CGCCTGCGCATCCCGCATCTTCTGCCCGCGGCCCGACAGATCGACCGCTTCGTCGAG 267
 |||||||
 Db 212 CGCCTGCGCATCCCGCATCTTCTGCCCGCGGCCCGACAGATCGACCGCTTCGTCGAG 271
 |||||||

QY 268 ATGTGACACGAGCGCAACGACGCGGGGAGAGGCTGTGGAGTGCACCTGTCTGGGCTTT 327
 |||||||
 Db 272 ATGTGACACGAGCGCAACGACGCGGGGAGAGGCTGTGGAGTGCACCTGTCTGGGCTTT 331
 |||||||

QY 328 GCGCGCACTGGCACCATGTGCTGCTGTACCTGGTGAAGAGCGGGGCTTGGCTGCAGGA 387
 |||||||
 Db 332 GCGCGCACTGGCACCATGTGCTGCTGTACCTGGTGAAGAGCGGGGCTTGGCTGCAGGA 391
 |||||||

QY 388 GATGCCATTCGTGAATTCGACGACACTACGACCGGCTCCATCGAGACCTATGACAGAG 447
 |||||||
 Db 392 GATGCCATTCGTGAATTCGACGACACTACGACCGGCTCCATCGAGACCTATGACAGAG 451
 |||||||

QY 448 AAGGCACTCTCCAGTCTCTACAGCGAAGAAATAGGGGCTTATGACCTTTACCAAG 507
 |||||||
 Db 452 AAGGCACTCTCCAGTCTCTACAGCGAAGAAATAGGGGCTTATGACCTTTACCAAG 511
 |||||||

QY 508 GCCCTCACTCCCTTCCCATGTGTGATGAGGGGCGAGAGTGAAGGAAAGTGGACTAAA 567
 |||||||
 Db 512 GCCCTCACTCCCTTCCCATGTGTGATGAGGGGCGAGAGTGAAGGAAAGTGGACTAAA 571
 |||||||

QY 568 GTATTAAACCTCTAGCTCCCATTTGCTGAAGACACTGAAGTACCCACCTTCAGAGCA 627
 |||||||
 Db 572 GTATTAAACCTCTAGCTCCCATTTGCTGAAGACACTGAAGTACCCACCTTCAGAGCA 631
 |||||||

QY 628 GGTTCGATTTGAAGGGAGGCTGTACTGCTTGTGTGAATTAATGAGTTTACGACCAA 687
 |||||||
 Db 632 GGTTCGATTTGAAGGGAGGCTGTACTGCTTGTGTGAATTAATGAGTTTACGACCAA 691
 |||||||

QY 688 AAAAAAAAAAAAAAAAAA 707
 |||||||
 Db 692 GAAAAAAAAAAAAAAAAA 711
 |||||||

RESULT 7
 AX482336 746 bp DNA linear PAT 16-AUG-2002
 LOCUS AX482336
 DEFINITION Sequence 5 from Patent WO02057460.
 ACCESSION AX482336
 VERSION AX482336.1 GI:22316926
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schlieven,G.,
 Suchard,S., Banas,D., Bassolino,D., Feder,J., Rysstek,S.,
 Mactee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramanathan,C.
 Polynucleotides encoding human phosphatases
 Patent: WO 02057460-A 5 25-JUL-2002;
 SQUIBB BRISTOL MYERS CO (US)

FEATURES

Location/Qualifiers
 1..746
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 178 a 221 c 218 g 129 t

ORIGIN

Query Match 95.3%; Score 673.6; DB 6; Length 746;
 Best Local Similarity 99.4%; Pred. No. 5.3e-111;
 Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 28 GCGATGGGGCTGCAGACCCCACTTCTCTGGGTGCTTCGCGGGCCGGCTGGCGGAGCTG 87
 |||||||
 Db 32 GCGATGGGGCTGCAGACCCCACTTCTCTGGGTGCTTCGCGGGCCGGCTGGCGGAGCTG 91
 |||||||

QY 88 GCGGTGCGCGGCTGCCCGCCCACTACAGAGTTCCTGTTGGACCTGGGGCTGGCGGAGCTG 147
 |||||||
 Db 92 GCGGTGCGCGGCTGCCCGCCCACTACAGAGTTCCTGTTGGACCTGGGGCTGGCGGAGCTG 151
 |||||||

QY 148 GTGTCCCTGACGAGCGGGGCGCCCTTCACAGCGACGTGCTCCCGGCTTCACCTGAC 207
 |||||||
 Db 152 GTGTCCCTGACGAGCGGGGCGCCCTTCACAGCGACGTGCTCCCGGCTTCACCTGAC 211
 |||||||

QY 208 CGCCTGCGCATCCCGCATCTTCTGCCCGCGGCCCGACAGATCGACCGCTTCGTCGAG 267
 |||||||
 Db 212 CGCCTGCGCATCCCGCATCTTCTGCCCGCGGCCCGACAGATCGACCGCTTCGTCGAG 271
 |||||||

QY 268 ATGTGACACGAGCGCAACGACGCGGGGAGAGGCTGTGGAGTGCACCTGTCTGGGCTTT 327
 |||||||
 Db 272 ATGTGACACGAGCGCAACGACGCGGGGAGAGGCTGTGGAGTGCACCTGTCTGGGCTTT 331
 |||||||

QY 328 GCGCGCACTGGCACCATGTGCTGCTGTACCTGGTGAAGAGCGGGGCTTGGCTGCAGGA 387
 |||||||
 Db 332 GCGCGCACTGGCACCATGTGCTGCTGTACCTGGTGAAGAGCGGGGCTTGGCTGCAGGA 391
 |||||||

QY 388 GATGCCATTCGTGAATTCGACGACACTACGACCGGCTCCATCGAGACCTATGACAGAG 447
 |||||||
 Db 392 GATGCCATTCGTGAATTCGACGACACTACGACCGGCTCCATCGAGACCTATGACAGAG 451
 |||||||

QY 448 AAGGCACTCTCCAGTCTCTACAGCGAAGAAATAGGGGCTTATGACCTTTACCAAG 507
 |||||||
 Db 452 AAGGCACTCTCCAGTCTCTACAGCGAAGAAATAGGGGCTTATGACCTTTACCAAG 511
 |||||||

QY 508 GCCCTCACTCCCTTCCCATGTGTGATGAGGGGCGAGAGTGAAGGAAAGTGGACTAAA 567
 |||||||
 Db 512 GCCCTCACTCCCTTCCCATGTGTGATGAGGGGCGAGAGTGAAGGAAAGTGGACTAAA 571
 |||||||

QY 568 GTATTAAACCTCTAGCTCCCATTTGCTGAAGACACTGAAGTACCCACCTTCAGAGCA 627
 |||||||
 Db 572 GTATTAAACCTCTAGCTCCCATTTGCTGAAGACACTGAAGTACCCACCTTCAGAGCA 631
 |||||||

QY 628 GGTTCGATTTGAAGGGAGGCTGTACTGCTTGTGTGAATTAATGAGTTTACGACCAA 687
 |||||||
 Db 632 GGTTCGATTTGAAGGGAGGCTGTACTGCTTGTGTGAATTAATGAGTTTACGACCAA 691
 |||||||

QY 688 AAAAAAAAAAAAAAAAAA 707
 |||||||
 Db 692 GAAAAAAAAAAAAAAAAA 711
 |||||||

RESULT 8
 AK000449 746 bp mRNA linear PRI 22-FEB-2000
 LOCUS AK000449
 DEFINITION Homo sapiens CDNA FLJ20442 fls, clone KAT04828.
 ACCESSION AK000449
 VERSION AK000449.1 GI:7020544
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (sites)

AUTHORS
Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H.,
Ota T., Suzuki Y., Ohashi M., Nishi T., Shibahara T., Tanaka T.,
Nakamura Y., Isogai T., and Sugano S.

JOURNAL
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 746)

REFERENCE
Sugano S., Suzuki Y., Ota T., Ohashi M., Nishi T., Isogai T.,
Shibahara T., Tanaka T., and Nakamura Y.

TITLE
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction: 5' - 6' 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source
1..746
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT04828"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_1db="KAT"
/note="cloning vector pME18SFL3"
35..487
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91172.1"
/db_xref="GI:7020545"
/translation="MGVQPPNFSNVLPRLGLALPRLPAHYOFLIDGVNHYSLTE
RGPNSDSCGLTARLRIPRCPAPDQIDRFQYIDENAREANVYKALFGRK
GRMACTIVERKGLAAGDALAIRLRRLRPLETEQEKAVFOYRRTK"

CDS
BASE COUNT 178 a 221 c 218 g 129 t
ORIGIN

Query Match 95.3%; Score 673.6; DB 9; Length 746;
Best Local Similarity 99.4%; Pred. No. 5.3e-111;
Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

28 GCGATGGCGCTGACAGCCCCCAACTTCTCTGCGTTCGCGGCGCGGCTGGCGGAGCTG 87
Db 32 GCGATGGCGCTGACAGCCCCCAACTTCTCTGCGTTCGCGGCGCGGCTGGCGGAGCTG 91

88 GCGCTGCGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 147
Db 92 GCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 151

148 GTGTCCTGAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 207
Db 152 GTGTCCTGAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 211

208 CGCGCTGCGCATCCCGACTTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 267
Db 212 CGCGCTGCGCATCCCGACTTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 271

268 ATGCTGAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 327
Db 272 ATGCTGAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 331

328 GCGCGCATGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
Db 332 GCGCGCATGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391

388 GATGCGCATGCTGAATCCGAGCTAGAGCCCGGCTCCATGAGACCTATGAGAGAGAG 447
Db 392 GATGCGCATGCTGAATCCGAGCTAGAGCCCGGCTCCATGAGACCTATGAGAGAGAG 451

448 AAGCAGCTCTTCAAGTTCTTACAGGAGAAATTAAGGGGCGCTTAGTACCTTCTACAG 507

Db 452 AAGCAGCTCTTCAAGTTCTTACAGGAGAAATTAAGGGGCGCTTAGTACCTTCTACAG 511

508 GCGCGCATGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
Db 512 GCGCGCATGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571

568 GTATTAAACCCCTCTAGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 627
Db 572 GTATTAAACCCCTCTAGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631

628 GGTCTGATGAGAGGAGGAGGCTTGTACTGCTTGTGTAATAATGATTTACGACCA 687
Db 632 GGTCTGATGAGAGGAGGAGGCTTGTACTGCTTGTGTAATAATGATTTACGACCA 691

688 AAAAAAAAAAAAAAAAAA 707
Db 692 GGAIAAAAAAAAAAAAAAAAAA 711

RESULT 9
AX398833 709 bp DNA linear PAT 27-MAY-2002
LOCUS AX398833
DEFINITION Sequence 4 from Patent WO0220747.
ACCESSION AX398833
VERSION AX398833.1 GI:21261345
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS Kossida S.
TITLE Regulation of human tyrosine phosphatase-like enzyme
JOURNAL Patent: WO 0220747-A 4 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source
1..709
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 137 a 233 c 206 g 133 t
ORIGIN

Query Match 95.0%; Score 671.6; DB 6; Length 709;
Best Local Similarity 98.6%; Pred. No. 1.2e-110;
Matches 688; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

1 TGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 13 TGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 72

61 GTGCTTCGGGCGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 73 GTGCTTCGGGCGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 132

121 CTGTTGAGACCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 133 CTGTTGAGACCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 192

181 GACAGCTCCCGGCGGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCG 240
Db 193 GACAGCTCCCGGCGGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCG 252

241 CCCGACAGATGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 253 CCCGACAGATGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312

301 GTGGAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 313 GTGGAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372

361 GTGAAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 373 GTGAAGGAGCGGGGCTTGGCTGACAGAGATGCATCTGCTGAATCCGACGATACGAGCC 432
QY 421 GGCTCATCGAGACCTATGACAGAGAGAAAGACAGTCTTCAGTTCTTACACGAAAGAA 480
Db 433 GGCTCATCGAGACCTATGACAGAGAGAAAGACAGTCTTCAGTTCTTACACGAAAGAA 492
QY 481 TAAAGGGGCTTATGATACCTTCTTACAGGCGCTCACTCCCTTCCTCCATGTTGCGATGG 540
Db 493 TAAAGGGGCTTATGATACCTTCTTACAGGCGCTCACTCCCTTCCTCCATGTTGCGATGG 552
QY 541 GCCAGAGATGAAGGAGTGTGACATTAAGTATTAACCTCTAGCTCCATTGGCTGAAGA 600
Db 553 GCCAGAGATGAAGGAGTGTGACATTAAGTATTAACCTCTAGCTCCATTGGCTGAAGA 612
QY 601 CACTGAAGTATGACCCCTGACAGGAGCTCTGATTAAGGAGGAGGCTTGTACTGCTTT 660
Db 613 CACTGAAGTATGACCCCTGACAGGAGCTCTGATTAAGGAGGAGGCTTGTACTGCTTT 671
QY 661 GTTGAATTAATGATTTTACGACCAAAAAAAAAAAAA 698
Db 672 GTTGAATTAATGATTTTACGACCAAAAAAAAAAAAA 709

RESULT 10

AX086042 694 bp DNA linear PAT 09-MAR-2001
LOCUS Sequence 39 from Patent WO0112819.
DEFINITION AX086042
ACCESSION AX086042.1 GI:13275852
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 694)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Plozman, G.D., Martinez, R., Whyte, D., Hill, R., Flanagan, P. and
JOURNAL Protein phosphatases and diagnosis and treatment of
Patent: WO 0112819-A 39 22-FEB-2001;
Sugen, Inc. (US)

FEATURES

source 1..694
Location/Qualifiers
BASE COUNT 128 a 221 c 216 g 129 t
ORIGIN

Query Match 93.0%; Score 657.4; DB 6; Length 694;
Best Local Similarity 99.8%; Pred. No. 4.2e-108;
Matches 658; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 28 GCATGAGGCGTGCAGCCCACTTCTCGTGGTGTTCGGGCGGCGGCGGAGCTG 87
Db 32 GCATGAGGCGTGCAGCCCACTTCTCGTGGTGTTCGGGCGGCGGCGGAGCTG 91
QY 88 GCCTGCGCGGCGTGCAGCCCACTTCTCGTGGTGTTCGGGCGGCGGAGCTG 147
Db 92 GCCTGCGCGGCGTGCAGCCCACTTCTCGTGGTGTTCGGGCGGCGGAGCTG 151
QY 148 GTTCTCTGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 207
Db 152 GTTCTCTGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 211
QY 208 CGCTGCGCGTGCAGCCCACTTCTCGTGGTGTTCGGGCGGCGGCGGCGGCGG 267
Db 212 CGCTGCGCGTGCAGCCCACTTCTCGTGGTGTTCGGGCGGCGGCGGCGGCGG 271
QY 268 ATGCTGAGAGGCGCAACGACGAGGAGAGGCTGTGGAGTGTGCTGTGGGCTTT 327
Db 272 ATGCTGAGAGGCGCAACGACGAGGAGAGGCTGTGGAGTGTGCTGTGGGCTTT 331
QY 328 GAGCGGAGTGTGAGCAATGCTGCTGTGTACCTGTGTGAAGAGCGGCGGCTTGGCTGAGGA 387

Db 332 GCGCGACATGGCACATCTGCTGCTTACTGTGGAAGAGCGGCGGCTTGGCTGAGGA 391
QY 388 GATGCCATTTGCTGAATTCGACGACACTACGACCCGCTTCATGAGACCTATGAGCAGAG 447
Db 392 GATGCCATTTGCTGAATTCGACGACACTACGACCCGCTTCATGAGACCTATGAGCAGAG 451
QY 448 AAAGCAGTCTTCCAGTCTTACACGAGAAATTAAGGCGCTTATGATACCTTTACACG 507
Db 452 AAAGCAGTCTTCCAGTCTTACACGAGAAATTAAGGCGCTTATGATACCTTTACACG 511
QY 508 GCCCTCACTCCCTTCCCATGTTGTGATGGGCGGAGATGAAGGGAAGTGAAGCTAAA 567
Db 512 GCCCTCACTCCCTTCCCATGTTGTGATGGGCGGAGATGAAGGGAAGTGAAGCTAAA 571
QY 568 GTATTAACCTCTAGCTCCCATTTGGCTGAAGACACTGAATACCCACCCCTGACGCA 627
Db 572 GTATTAACCTCTAGCTCCCATTTGGCTGAAGACACTGAATACCCACCCCTGACGCA 631
QY 628 GGTCTGATTAAGGAGGCGGCTTACTGCTTTGTGAATTAATGATTTACGACCA 686
Db 632 GGTCTGATTAAGGAGGCGGCTTACTGCTTTGTGAATTAATGATTTACGACCA 690

RESULT 11

AX398836 602 bp DNA linear PAT 27-MAY-2002
LOCUS Sequence 7 from Patent WO0220747.
DEFINITION AX398836
ACCESSION AX398836
VERSION AX398836.1 GI:21261348
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Kossida, S.
TITLE Regulation of human tyrosine phosphatase-like enzyme
JOURNAL Patent: WO 0220747-A 7 14-MAR-2002;
Bayer Aktiengesellschaft (DE)

FEATURES

source 1..602
Location/Qualifiers
BASE COUNT 109 a 202 c 179 g 110 t 2 others
ORIGIN

Query Match 82.6%; Score 583.8; DB 6; Length 602;
Best Local Similarity 99.3%; Pred. No. 6.5e-95;
Matches 596; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 42 GCGCCCACTTCTCTGCTGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCT 101
Db 3 GCGCCCACTTCTCTGCTGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCT 62
QY 102 CCGCGCCACTTACAGTCTCTGTTGACGCTGGGCGTGGCGGCGGCGGCGGCGG 161
Db 63 CCGCGCCACTTACAGTCTCTGTTGACGCTGGGCGTGGCGGCGGCGGCGGCGG 122
QY 162 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 221
Db 123 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
QY 222 CGACTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 281
Db 183 CGACTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
QY 282 CAACGACGAGGAGAGGCTGTGGAGTGTGCTGTGGGCTTGGCGGCGGCGGCGG 341
Db 243 CAACGACGAGGAGAGGCTGTGGAGTGTGCTGTGGGCTTGGCGGCGGCGGCGG 302
QY 342 CATGCTGCTGCTTACTGTGTGAAGGAGGCGGCTTGGCTGAGAGATGATGCTG 400

Db	303	CATCTGGCCGTGTACCTGGTGAANGAGCGNGCTTGCTGCAGAGATGCCATTGGCG	362
QY	401	AAATCCGACGACTACGACCCGGCTCCATCGAGACTATGACGAGAGAAAGCATGCTTCC	460
Db	363	AAATCCGACGACTACGACCCGGCTCCATCGAGACTATGACGAGAGAAAGCATGCTTCC	422
QY	461	AGTTCTACACGAGCAAGAAATTAAGGGGCTTAGTACCTCTTACAGAGCCCTCATCTCCC	520
Db	423	AGTTCTACACGAGCAAGAAATTAAGGGGCTTAGTACCTCTTACAGAGCCCTCATCTCCC	482
QY	521	TTTCCCATGTTGTGATGAGGGGCCACAGATGAAGGAACTGGACTTAAAGTATTTAAACCTTC	580
Db	483	TTTCCCATGTTGTGATGAGGGGCCACAGATGAAGGAACTGGACTTAAAGTATTTAAACCTTC	542
QY	581	TAGTCCCATGTCGCGTGAAGACATCAAGATTAAGCCCACTCTGAGGCAAGCTCTGATTGAA	640
Db	543	TAGTCCCATGTCGCGTGAAGACATCAAGTAACTAGCCCACTCTGAGGCAAGCTCTGATTGAA	602

RESULT 12	LOCUS	DEFINITION	559 bp	mrna	linear	PRI 29-AUG-1998
HUMZE12H05	HUMZE12H05	Homo sapiens full length insert cDNA clone ZEL12H05.				
		reverse52				

SOURCE ORGANISM

REFERENCE AUTHORS

TITLE	Full Clone Sequencing of the Longest Available Member from Each
JOURNAL	unigene Cluster
REFERENCE	2 (bases 1 to 559)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (24-MUG-1998)
COMMENT	Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
	SUBMITTED BY:

NOTE: This sequence represents the full insert of this cDNA. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived. We have tried to obtain double-stranded, or double chemistry sequence across the entire clone, but potentially, there are areas in the sequence where this level of coverage was not achieved. Nevertheless, we are confident of the accuracy of this sequence as all regions of low quality, as defined by PhRAP (P. Green, in preparation), were visually inspected and edited accordingly. The consensus quality values for this sequence have been submitted separately.

SIMILARITY INFORMATION:

similar to Clona intestinalis protein Z83760 (PID:g1764017) COS41.7

The location of this clone is unknown.

FEATURES

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

BASE COUNT	121 a	169 c	162 g	107 t
misc_feature	/clone="IMAGE:358809" /clone_id="Soares_fetal_heart_NMH19W"			
misc_feature	54. .339 /note="similar to Clona intestinalis protein 283760 (PID:g1764017)"			
misc_feature	61. .330 /note="similar to Homo sapiens protein AF064105 (PID:g3136334)"			
misc_feature	85. .333 /note="similar to Pyrococcus horikoshii protein AB009524 (PID:g3132001)"			
misc_feature	91. .297 /note="similar to Caenorhabditis elegans protein U28739 (PID:g2731378)"			
misc_feature	91. .297 /note="similar to Caenorhabditis elegans protein U28739 (PID:g2731379)"			

Query Match	78.6%;	Score 555.8;	DB 9;	Length 559;
Best Local Similarity	99.6%;	Pred. No. 6.7e-90;		
Matches 557; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

DY 133 GGGCTGCGGCACCTGGGTCCCTGACGGAGCGCGGGGCCCTCACAAGGCACACTGCCCC 192
|||||
|||
DQ 1 GGGCTGCGGCACCTGGGTCCCTGACGGAGCGCGGGGCCCTCACAAGGCACACTGCCCC 60

193 GGGCTCACCCTGCACCGCCCTGGGCATCCCCGACTTCTGCCCCGGCCCCGACAGATC 252
 Db 61 GGGCTCACCCTGCACCGCCCTGGGCATCCCCGACTTCTGCCCCGGCCCCGACAGATC 120

121 GACCCTTCGTCAGATCGTGGACGAGGCCAACGCACGGGGAGAGCGCTGTGGAGTGCAC 180

Db 181 TGTGCTGGGCTTTGGCCGCACTGGCACCAGATGCGCTGTTACCTGGTGAAGGAGCGG 240

Db	241	GGCTTGGCTGCAGAGATGCCATTGCTGAAATCCGACGACTACGACCCGGCTTCATCGAG	300
07	433	ACCTATGAGCAGGAGAAAGCAGTCTTCCAGTTCCTACGCGGAAGAAATAAGGGGCTTAA	492

Db	301	ACCTATGAGCAGGAGAAAGCAGTCTTCCAGTTCTACCAAGCGAAGAAATTAAGGGCCCTTA	360
Qy	493	GTACCTTCTACCAAGCCCTCACTCCCTTCCCATGTTGTGATGGGCCAGAGATGA	552

Db 361 GTACCCCTTACACAGCCCTCACTCCCTTCCCCATGTGTGCATGGGGCCAGAGATGAA 420

DY 553 GGGAACTGGAATAAGTATTAAACCTCTAGCTCCATTGGCTGAAGACACTGAAGTACG 612

CCACCCCTGCAGGCGACGCTCGATTGTAAGGGAGGCTTGACGCTTTCTTGAATAATG 613
|||||

QY 673 AGTTTACGAACCAAAA 691
|||||
541 TCGGGGACCAATTTTTC 560

RESULT 13
AX398838

Accession	Sequence	from Patent	MO0220747.
AX398838			
AX398838.1	GI:21261350		

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Kossida, S.
TITLE Regulation of human tyrosine phosphatase-like enzyme
JOURNAL Patent: WO 0220747-A 9 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1.353
Location/Qualifiers
BASE COUNT 96 a 186 c 168 g 103 t
ORIGIN
Query Match 71.9%; Score 508.6; DB 6; Length 553;
Best Local Similarity 98.6%; Pred. No. 1.9e-81;
Matches 545; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
QY 29 CGATGGCGGTGACCCGCCCAACTTCTCTGGTGTCCGGGCGGCTGGGAGACTGG 88
DB 1 CGATGGCGGTGACCCGCCCAACTTCTCTGGTGTCCGGGCGGCTGGGAGACTGG 60
QY 89 CGGTGGCGGGGCTCCCGCCCACTACAGTCTCTGTGGACCTGGGCGTGGGACCTGG 148
DB 61 CGCTGGCGGGGCTCCCGCCCACTACAGTCTCTGTGGACCTGGGCGTGGGAGACTGG 120
QY 149 TGTGCTGAC--GGAGCGCGGGGCGGCGGCTCACAGGACAGTGGCGGCGCTCACCTGCA 206
DB 121 TGTGCTGACGTGGAGCGTGGGCGGCGGCGGCTCACAGGACAGTGGCGGCGCTCACCTGCA 180
QY 207 CCGGCTGGCATGCC--GACTTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCTGGC 265
DB 181 CCGGCTGGCATGCCCGCTGACTTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCTGGC 240
QY 266 AGATCGGAGCA--GGCAACGCGCGGGGAGAGGCTGGGAGGACGCTGTGCTGGGC 324
DB 241 AGATCGGAGCAAGGCGCAAGGCGGAGAGGCTGTGGAGTGCATGTGCTGGGC 300
QY 325 TTTGGCGCACTGGCACATGCTGGCTGTACTGTGTAAGAGCGGGGCTTGGTGA 384
DB 301 TTTGGCGCACTGGCACATGCTGGCTGTACTGTGTAAGAGCGGGGCTTGGTGA 360
QY 385 GGAGATGCCATGTGCTGAATCCGACGACTAGACCGCGGCTCATGAGACCTATGACAG 444
DB 361 GGAGATGCCATGTGCTGAATCCGACGACTAGACCGCGGCTCATGAGACCTATGACAG 420
QY 445 GAGAAAGCACTCTTCCAGTCTACAGGAGCAAGAAATTAAGGGGCTTAACTTCTAC 504
DB 421 GAGAAAGCACTCTTCCAGTCTACAGGAGCAAGAAATTAAGGGGCTTAACTTCTAC 480
QY 505 CAGGCGCTCACTCCCTTCCCATGTGTGATGGGCGGCGAGATGAAGGAAGTGGACT 564
DB 481 CAGGCGCTCACTCCCTTCCCATGTGTGATGGGCGGCGAGATGAAGGAAGTGGACT 540
QY 565 AAAGTATTAAAC 577
DB 541 AAAGTATTAAAC 553
RESULT 14
AX398834 493 bp DNA linear PAT 27-MAY-2002
LOCUS AX398834
DEFINITION Sequence 5 from Patent WO0220747.
ACCESSION AX398834
VERSION AX398834.1 GI:21261346
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Kossida, S.

TITLE Regulation of human tyrosine phosphatase-like enzyme
JOURNAL Patent: WO 0220747-A 5 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1.493
Location/Qualifiers
BASE COUNT 78 a 176 c 149 g 90 t
ORIGIN
Query Match 68.8%; Score 486.2; DB 6; Length 493;
Best Local Similarity 99.4%; Pred. No. 2e-77;
Matches 488; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 42 GCCCCCACTTCTCTGAGTCTTCCGGGCGGCGGAGTGGGCGGCTGGCGGCT 101
DB 3 GGCGCCCACTTCTCTGAGTCTTCCGGGCGGCGGAGTGGGCGGCTGGCGGCT 62
QY 102 CCGCGCCCACTTCCAGTCTCTGTGACCTGGGCGTGGGCGGCTGGTGTCCGACGGA 161
DB 63 CCGCGCCCACTTCCAGTCTCTGTGACCTGGGCGTGGGCGGCTGGTGTCCGACGGA 122
QY 162 GCGGGGCGGCGGCTTACAGCGACAGTCCCGGCGCTTACCTGACCGCTGGCATGCC 221
DB 123 GCGGGGCGGCGGCTTACAGCGACAGTCCCGGCGCTTACCTGACCGCTGGCATGCC 182
QY 222 GCACTTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 281
DB 183 GCACTTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
QY 282 CAACGACGGGGAGAGGCTGTGGAGTGCATGTGCTGTGGGCTTGGCGGCGGCGGCGG 341
DB 243 CAACGACGGGGAGAGGCTGTGGAGTGCATGTGCTGTGGGCTTGGCGGCGGCGGCGG 302
QY 342 CATGCTGGCTGTACTGTGTAAGAGCGGGGCTTGGCTGACGAGATGCTATGCTGA 401
DB 303 CATGCTGGCTGTACTGTGTAAGAGCGGGGCTTGGCTGACGAGATGCTATGCTGA 362
QY 402 AATCGACGACTAGACCGGCGGCTCATGAGACTATGAGAGGAAAGAGAGTGTCCA 461
DB 363 AATCGACGACTAGACCGGCGGCTCATGAGACTATGAGAGGAAAGAGAGTGTCCA 422
QY 462 GTTCTACGACGAAAGAAATTAAGGGGCTTAACTTCTACAGGCGGCTCACTGCCCT 521
DB 423 GTTCTACGACGAAAGAAATTAAGGGGCTTAACTTCTACAGGCGGCTCACTGCCCT 482
QY 522 TCCCATGTTG 532
DB 483 TCCCATGTTG 493
RESULT 15
AX398839 448 bp DNA linear PAT 27-MAY-2002
LOCUS AX398839
DEFINITION Sequence 10 from Patent WO0220747.
ACCESSION AX398839
VERSION AX398839.1 GI:21261351
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Kossida, S.
TITLE Regulation of human tyrosine phosphatase-like enzyme
JOURNAL Patent: WO 0220747-A 10 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1.448
Location/Qualifiers
BASE COUNT 93 a 132 c 120 g 103 t
ORIGIN

Query Match 63.1%; Score 446; DB 6; Length 448;
Best Local Similarity 100.0%; Pred. No. 3.1e-70;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 CCGGACGATGACCGCGCTTCGTGAGATCGTGACGAGGCCAACGACGCGGAGAGGCT 300
|||||
DB 448 CCGGACGATGACCGCGCTTCGTGAGATCGTGACGAGGCCAACGACGCGGAGAGGCT 389
|||||
QY 301 GTGGAGTGCATGTGCTCTGCGCTTTGGCCGCTGCGACCATGTCGCTGTTACCTG 360
|||||
DB 388 GTGGAGTGCATGTGCTCTGCGCTTTGGCCGCTGCGACCATGTCGCTGTTACCTG 329
|||||
QY 361 GTGAAGAGCGGGGCTTGCGTGACGAGATGCCATTCGTAATCGACGACTACGACC 420
|||||
DB 328 GTGAAGAGCGGGGCTTGCGTGACGAGATGCCATTCGTAATCGACGACTACGACC 269
|||||
QY 421 GGCTCCATCGAGACCTATGAGAGAGAAAGAGTCTTCAGTCTTACGACGAAAGAAA 480
|||||
DB 268 GGCTCCATCGAGACCTATGAGAGAGAAAGAGTCTTCAGTCTTACGACGAAAGAAA 209
|||||
QY 481 TAAGGGGCTTAGTACCTTCTACGAGGCGCTCACTCCGCTTCCCATGTTGCGATGGG 540
|||||
DB 208 TAAGGGGCTTAGTACCTTCTACGAGGCGCTCACTCCGCTTCCCATGTTGCGATGGG 149
|||||
QY 541 GCCAGAGATGAAAGGAGTGAATTAAGTAAACCTCTAGTCCCATTTGGCTGAAGA 600
|||||
DB 148 GCCAGAGATGAAAGGAGTGAATTAAGTAAACCTCTAGTCCCATTTGGCTGAAGA 89
|||||
QY 601 CACTGAATAGCCCAACCCCTGAGGAGGCTGCTGATGAAGGGAGGCTTGTACTGCTTT 660
|||||
DB 88 CACTGAATAGCCCAACCCCTGAGGAGGCTGCTGATGAAGGGAGGCTTGTACTGCTTT 29
|||||
QY 661 GTTGAATTAATGAGTTTACGAAACA 686
|||||
DB 28 GTTGAATTAATGAGTTTACGAAACA 3
|||||

Search completed: November 4, 2002, 08:49:28
Job time : 2789 secs

Query Match	100.0%	Score 707	DB 22	Length 707
Best Local Similarity	100.0%	Pred. No. 1.9e-138		
Matches 707	Conservative 0	Mismatches 0	Indels 0	Gaps 0
<p>Sequence 707 BP; 146 A; 224 C; 202 G; 135 T; 0 other;</p>				
YY	1	TCACCCGGCTGCTCTGTGCGCCCTTCCACAGGATGGGGGTGACAGCCGCCCACTTCTCTCG	60	
DB	1	TCACCCGGCTGCTCTGTGCGCCCTTCCACAGGATGGGGGTGACAGCCGCCCACTTCTCTCG	60	
YY	61	GTGCTTCCGGGGCGGCTGTGGGAGTGGGGGTGCGCGCGGCTCCCGCCACATACAGTTC	120	
DB	61	GTGCTTCCGGGGCGGCTGTGGGAGTGGGGGTGCGCGCGGCTCCCGCCACATACAGTTC	120	
YY	121	CTGTGGAGACTGGGGGTGGCGGACCTGTGTCTTCACGAGACGGGGGCCCTTCACAC	180	
DB	121	CTGTGGAGACTGGGGGTGGCGGACCTGTGTCTTCACGAGACGGGGGCCCTTCACAC	180	
YY	181	GACACGTGCCCCGGGCGCTCACCCCTGCACCGCCTTGGCATCCCGGACTTTCGCCCGCGCC	240	
DB	181	GACACGTGCCCCGGGCGCTCACCCCTGCACCGCCTTGGCATCCCGGACTTTCGCCCGCGCC	240	
YY	241	CCGACACGATGAGCGCGCTTGTGAGATCTGTGAGAGAGAGCCACGACAGGGGAGAGGCT	300	
DB	241	CCGACACGATGAGCGCGCTTGTGAGATCTGTGAGAGAGAGCCACGACAGGGGAGAGGCT	300	
YY	301	GTGGAGTGCACCTGTGCTGTGGGCTTGGGCGCGACCTGGGACCATGTGGCTGTTCAC	360	
DB	301	GTGGAGTGCACCTGTGCTGTGGGCTTGGGCGCGACCTGGGACCATGTGGCTGTTCAC	360	
YY	361	GTGAAGGAGCGGGGCTTGGCTGCAGAGATGTCATTTGCTGAATCCGACGACTACGACCC	420	
DB	361	GTGAAGGAGCGGGGCTTGGCTGCAGAGATGTCATTTGCTGAATCCGACGACTACGACCC	420	
YY	421	GGCTCATCTGAGACCTATAGACAGAGAAAGAGTCTTCACATTTACACAGGAA	480	
DB	421	GGCTCATCTGAGACCTATAGACAGAGAAAGAGTCTTCACATTTACACAGGAA	480	
YY	481	TAAAGGGGCTTAGTACCTTTCACAGGCGCCCTACCTCCCTTCCCATTTGTCATGGG	540	
DB	481	TAAAGGGGCTTAGTACCTTTCACAGGCGCCCTACCTCCCTTCCCATTTGTCATGGG	540	
YY	541	GCCACAGATGAAAGGAGAGTGAATTAAGTATTAACCTCTAGTCTCCATTGGCTGAAGA	600	
DB	541	GCCACAGATGAAAGGAGAGTGAATTAAGTATTAACCTCTAGTCTCCATTGGCTGAAGA	600	
YY	601	CACGTAATAGCCGACCCCTGACGAGAGTCTGTATTAAGGGGAGGCTTGTACTGCTTT	660	
DB	601	CACGTAATAGCCGACCCCTGACGAGAGTCTGTATTAAGGGGAGGCTTGTACTGCTTT	660	
YY	661	GTTGAATTAATGAGTTTACGAAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	707	
DB	661	GTTGAATTAATGAGTTTACGAAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	707	

[illegible]

```
QY 61 GTGCTTCGGGCGGCGGAGCTGGGCTGCTGCCGCGGCTCCCGCCACTACCACTTC 120
   |||
   |||
   |||
Db 134 GTGCTTCGGGCGGCGGAGCTGGGCTGCTGCCGCGGCTCCCGCCACTACCACTTC 193
   |||
   |||
   |||
QY 121 CTTTGGAGCTGGGCTGGCGGACCTGGTGTCCCTGACGAGCGGGGCCCTCACAGC 180
   |||
   |||
   |||
Db 194 CTTTGGAGCTGGGCTGGCGGACCTGGTGTCCCTGACGAGCGGGGCCCTCACAGC 253
   |||
   |||
   |||
QY 181 GAGAGTGGCGGCGGCTGACGACCGGCTGGGCTGCCGCGGCTTCGCGCGGCGG 240
   |||
   |||
   |||
Db 254 GAGAGTGGCGGCGGCTGACGACCGGCTGGGCTGCCGCGGCTTCGCGCGGCGG 313
   |||
   |||
   |||
QY 241 CCCGACGATGACCGCTTGTGAGAGTGGGACGAGCGGCGGAGAGGCT 300
   |||
   |||
   |||
Db 314 CCCGACGATGACCGCTTGTGAGAGTGGGACGAGCGGCGGAGAGGCT 373
   |||
   |||
   |||
QY 301 GTGGAGTGCATGTGCTTGGGCTTGGCGGACGCTGGACATGCTGGCTGTACCTG 360
   |||
   |||
   |||
Db 374 GTGGAGTGCATGTGCTTGGGCTTGGCGGACGCTGGACATGCTGGCTGTACCTG 433
   |||
   |||
   |||
QY 361 GTGAGGAGCGGGGCTTGGGCTGGAGAGATGCCATGCTGAAATCCGAGCGACTGAG 420
   |||
   |||
   |||
Db 434 GTGAGGAGCGGGGCTTGGGCTGGAGAGATGCCATGCTGAAATCCGAGCGACTGAG 493
   |||
   |||
   |||
QY 421 GGCTCCATCGAGACTATGAGCAGAGAGAAAGCACTTCCAGTCTACAGCGAAGAAA 480
   |||
   |||
   |||
Db 494 GGCTCCATCGAGACTATGAGCAGAGAGAAAGCACTTCCAGTCTACAGCGAAGAAA 553
   |||
   |||
   |||
QY 481 TAAGGGGCTTATGACCTTCTACAGAGCCCTACCTCCCTCCCATGTTGTGATGGG 540
   |||
   |||
   |||
Db 554 TAAGGGGCTTATGACCTTCTACAGAGCCCTACCTCCCTCCCATGTTGTGATGGG 613
   |||
   |||
   |||
QY 541 GCCAGAGTGAAGGAGTGAAGTAAAGTAAACCTAGTCCCATGCTGAGAG 600
   |||
   |||
   |||
Db 614 GCCAGAGTGAAGGAGTGAAGTAAAGTAAACCTAGTCCCATGCTGAGAG 673
   |||
   |||
   |||
QY 601 CACTGAAGTAGCCACCTGTCAGAGGAGTCTGATTAAGAGGAGGCTGTGCTT 660
   |||
   |||
   |||
Db 674 CACTGAAGTAGCCACCTGTCAGAGGAGTCTGATTAAGAGGAGGCTGTGCTT 733
   |||
   |||
   |||
QY 661 GTTGAATTAATGATTTTACGAACCAAAAAA 703
   |||
   |||
   |||
Db 734 GTTGAATTAATGATTTTACGAACCAAAAAA 776
   |||
   |||
   |||
RESULT 3
ABLA0814
ID ABLA0814 standard; DNA; 746 BP.
XX
AC ABLA0814;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human tyrosine phosphatase-like enzyme related DNA sequence.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
   cerebroprotective; immunomodulator; antiparkinsonian; anticonvulsant;
   antihypertensive; neuroprotective; nootropic; neuroleptic; anti-HIV;
   antihypertensive; hypotensive; vulnerability; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200220747-A2.
XX
PD 14-MAR-2002.
XX
PF 05-SEP-2001; 2001WO-EPI0205.
XX
PR 11-SEP-2000; 2000US-231568P.
XX
PR 06-DEC-2000; 2000US-251403P.
XX
PA (FARB ) BAYER AG.
XX
```

```
PI Kossida S;
XX
DR WPI: 2002-339803/37.
XX
PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
   PT which are useful for preventing, treating diabetes, obesity, cancer,
   PT cardiovascular and pulmonary diseases
XX
PS Disclosure; Fig 11; 117pp; English.
XX
CC The invention relates to a purified human tyrosine phosphatase-like
   CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
   CC methodology. The tyrosine phosphatase-like enzyme and encoding
   CC polynucleotides are useful for screening for modulators which are used
   CC for treating a tyrosine phosphatase-like enzyme dysfunction related
   CC disease such as a central nervous system (CNS) disorder, diabetes,
   CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
   CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
   CC other eating disorder such as bulimia. CNS disorders include brain
   CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
   CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
   CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
   CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
   CC and cardiovascular diseases include myocardial infarction, ischaemic
   CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
   CC vascular diseases and peripheral vascular diseases. The enzyme is useful
   CC in diagnostic assays for detecting diseases and abnormalities or
   CC susceptibility to diseases or abnormalities related to the presence of
   CC mutations in the encoding nucleic acid sequences. The present sequence
   CC represents a human tyrosine phosphatase-like enzyme related DNA.
   CC
SQ Sequence 746 BP; 178 A; 221 C; 218 G; 129 T; 0 other;
XX
Query Match 95.3%; Score 673.6; DB 24; Length 746;
Best Local Similarity 99.4%; Pred. No. 1.8e-131;
Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 28 GCGATGGGCGTGCAGACCCCGCACTTCTGAGTCTTCCGGGCGGCTGGGAGCTG 87
   |||
   |||
   |||
Db 32 GCGATGGGCGTGCAGACCCCGCACTTCTGAGTCTTCCGGGCGGCTGGGAGCTG 91
   |||
   |||
   |||
QY 88 GCGCTGGGCGGCTGCCCGGCACTACAGTCTGTTGGAGCTGGGCGTGGGACCTG 147
   |||
   |||
   |||
Db 92 GCGCTGGGCGGCTGCCCGGCACTACAGTCTGTTGGAGCTGGGCGTGGGACCTG 151
   |||
   |||
   |||
QY 148 GTGTCCCTGAGGAGCGGCGGCGGCTTCACAGGAGAGCTGCCGCGGCTTCCCTGAC 207
   |||
   |||
   |||
Db 152 GTGTCCCTGAGGAGCGGCGGCGGCTTCACAGGAGAGCTGCCGCGGCTTCCCTGAC 211
   |||
   |||
   |||
QY 208 GCGCTGGGCGATCCCGGATCTGCGCCCGGCGGCGGCGGCGGCGGCTTCCCTGAC 267
   |||
   |||
   |||
Db 212 GCGCTGGGCGATCCCGGATCTGCGCCCGGCGGCGGCGGCGGCGGCTTCCCTGAC 271
   |||
   |||
   |||
QY 268 ATCTGTCAGAGGCGGCGGAGAGGCTGTGGAGTGCATGTGCTGTGGGCTT 327
   |||
   |||
   |||
Db 272 ATCTGTCAGAGGCGGCGGAGAGGCTGTGGAGTGCATGTGCTGTGGGCTT 331
   |||
   |||
   |||
QY 328 GCGCGCACTGGCAGCATGCTGGCTGTACTGTGGTGAAGAGCGGAGGCTTGGCTGACAGA 387
   |||
   |||
   |||
Db 332 GCGCGCACTGGCAGCATGCTGGCTGTACTGTGGTGAAGAGCGGAGGCTTGGCTGACAGA 391
   |||
   |||
   |||
QY 388 GATGCAATGTCTGAATCCAGAGCTACGACCGGCTCCATGAGACCTATGAGAGAG 447
   |||
   |||
   |||
Db 392 GATGCAATGTCTGAATCCAGAGCTACGACCGGCTCCATGAGACCTATGAGAGAG 451
   |||
   |||
   |||
QY 448 AAGAGAGTCTTCCAGTGTACAGAGGAGAAATTAAGGGGCTTATGATCCTTTACAG 507
   |||
   |||
   |||
Db 452 AAGAGAGTCTTCCAGTGTACAGAGGAGAAATTAAGGGGCTTATGATCCTTTACAG 511
   |||
   |||
   |||
QY 508 GCCCTCACTCCCTTCCCATGTTGTGATGGGCGGAGAGTGAAGGAGTGAAGTGAAG 567
   |||
   |||
   |||
Db 512 GCCCTCACTCCCTTCCCATGTTGTGATGGGCGGAGAGTGAAGGAGTGAAGTGAAG 571
   |||
   |||
   |||
```

QY 568 GATTTAAACCTTAGCTCCATGGCTGAGACACTGAGTACCCACCCCTGCAGGCA 627
DB 572 GATTTAAACCTTAGCTCCATGGCTGAGACACTGAGTACCCACCCCTGCAGGCA 631
QY 628 GGTCTGATTGAGGAGGAGGCTTGTACTGCTTGTGATTAATGAGTTTACGAAACCA 687
DB 632 GGTCTGATTGAGGAGGAGGCTTGTACTGCTTGTGATTAATGAGTTTACGAAACCA 691
QY 688 AAAAAAAAAAAAAAAAAAAAAA 707
DB 692 GGAIAAAAAAAAAAAAAAAAAA 711

RESULT 4
ABLA0815
ID ABLA0815 standard; cDNA; 746 BP.
AC ABLA0815;
XX
XX 03-JUL-2002 (first entry)
DE Human tyrosine phosphatase-like enzyme encoding cDNA.
XX
XX Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antihypertensive; hypotensive; vulnerary; gene therapy; human; gene; ss.
OS Homo sapiens.
XX
XX W0200220747-72.
XX
XX 14-MAR-2002.
XX
XX 05-SEP-2001; 2001WO-EP10205.
XX
XX 11-SEP-2000; 2000US-231568P.
XX 06-DEC-2000; 2000US-251403P.
XX
XX (FARB) BAYER AG.
XX
XX Kossida S;
XX
XX WPI; 2002-339803/37.
XX
XX New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases
XX
XX Claim 19; Page 117; 117pp; English.
XX
XX The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. The present sequence
CC represents the human tyrosine phosphatase-like enzyme polypeptide
XX encoding cDNA.
XX

SO Sequence 746 BP; 178 A; 221 C; 218 G; 129 T; 0 other;
Query Match 95.3%; Score 673.6; DB 24; Length 746;
Best Local Similarity 99.4%; Pred. No. 1.8e-131;
Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 28 GCGATGGGCGGCGACGCCCCCAACTTCCTGCGGCTTCCGGGCGGCGGAGCTG 87
DB 32 GCGATGGGCGGCGACGCCCCCAACTTCCTGCGGCTTCCGGGCGGCGGAGCTG 91
QY 88 GCGCTGCGGCGGCTTCCCGCCACTACAGTTCTGTTGACCTGGCGTGGCAGCTG 147
DB 92 GCGCTGCGGCGGCTTCCCGCCACTACAGTTCTGTTGACCTGGCGTGGCAGCTG 151
QY 148 GGTGCTGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 207
DB 152 GTGTCTGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 211
QY 208 GCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 267
DB 212 GCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 271
QY 268 ATCTGAGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 327
DB 272 ATCTGAGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 331
QY 328 GCGCGCACTGGCAGCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 387
DB 332 GCGCGCACTGGCAGCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 391
QY 388 GATGCGATTTGCTGAATCCGACGAGTACGACCCGCGCTCATGAGCTATGAGCAG 447
DB 392 GATGCGATTTGCTGAATCCGACGAGTACGACCCGCGCTCATGAGCTATGAGCAG 451
QY 448 AAAGCAGCTTCCAGTTTACGAGGAAAGAAATGAGGCGCTTACCTTACAG 507
DB 452 AAAGCAGCTTCCAGTTTACGAGGAAAGAAATGAGGCGCTTACCTTACAG 511
QY 508 GCGCTGAGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 567
DB 512 GCGCTGAGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 571
QY 568 GATTTAAACCTTAGCTCCATGGCTGAGACACTGAGTACCCACCCCTGCAGGCA 627
DB 572 GATTTAAACCTTAGCTCCATGGCTGAGACACTGAGTACCCACCCCTGCAGGCA 631
QY 628 GGTCTGATTGAGGAGGAGGCTTGTACTGCTTGTGATTAATGAGTTTACGAAACCA 687
DB 632 GGTCTGATTGAGGAGGAGGCTTGTACTGCTTGTGATTAATGAGTTTACGAAACCA 691
QY 688 AAAAAAAAAAAAAAAAAAAAAA 707
DB 692 GGAIAAAAAAAAAAAAAAAAAA 711

RESULT 5
ABLA0807
ID ABLA0807 standard; DNA; 709 BP.
AC ABLA0807;
XX
XX 03-JUL-2002 (first entry)
XX
XX Human tyrosine phosphatase-like enzyme related EST sequence #1.
XX
XX Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antihypertensive; hypotensive; vulnerary; gene therapy; human; gene;
XX expressed sequence tag; EST; ds.
XX
XX Homo sapiens.
XX

CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, Papillon-Lévy
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
CC Björnstad syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamartomas.

SD Sequence 694 BP; 128 A; 221 C; 216 G; 129 T; 0 other;

Query Match	93.08;	Score 657.4;	DB 22;	Length 694;
Best Local Similarity	99.88;	Pred. No. 4.3e-128;		
Matches 658; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	28	GCAGAGGGCGAGACGCCCCCAACTTCACCTCCAGGGAGCTTCCGGGGCCGGCTGGCCGAGCTG	87
DB	32	GCAGATGGGCGTGGACAGCCCCCAACTTCTCTGGGGGCTTCCGGGGCCGGCTGGCCGAGCTG	91
QY	88	GCAGCTGCAGCGAGCTCCCGCGCCACTACCACTTCCTGTGTGGACCTGGGGGTGCGGCACCTG	147
DB	92	GCAGCTGCAGCGAGCTCCCGCGCCACTACCACTTCCTGTGTGGACCTGGGGGTGCGGCACCTG	151
QY	148	GTGTTCCTTGACGAGAGCGGGGGCCCCCTACACGACGACCTGCCCGGCTCACCCTGCAC	207
DB	152	GTGTTCCTTGACGAGAGCGGGGGCCCCCTCACACGACGACCTGCCCGGCTCACCCTGCAC	211
QY	208	CGCCCTGGGCATCCCCGAACTTCTGGCCCGCGGGCCCCCGACACGATCGACCGCTTGTGACG	267
DB	212	CGCCCTGGGCATCCCCGAACTTCTGGCCCGCGGGCCCCCGACACGATCGACCGCTTGTGACG	271
QY	268	ATCTGTGACGAGGGCCAAACGCACGGGGAGAGGGCTGTGGAGTGCACCTGTGCTCGGGCTT	327
DB	272	ATCTGTGACGAGGGCCAAACGCACGGGGAGAGGGCTGTGGAGTGCACCTGTGCTCGGGCTT	331
QY	328	GGCCGCACTGGCACACAGCTGGCTGGCTTACCTGGTGAAGAGAGCGGGGCTTGGCTGCGAGA	387
DB	332	GGCCGCACTGGCACACAGCTGGCTGGCTTACCTGGTGAAGAGAGCGGGGCTTGGCTGCGAGA	391
QY	388	GATGCCATTGGCTGAATTCGACGACACTACGACCCCGGCTCTACACGAGACCTATGAGACAGAG	447
DB	392	GATGCCATTGGCTGAATTCGACGACACTACGACCCCGGCTCTACACGAGACCTATGAGACAGAG	451
QY	448	AAACAGAGCTTCCAGGTTCTACACAGCAAGCAAAATAAGGGGCTTATGATACCTTTACCAG	507
DB	452	AAACAGAGCTTCCAGGTTCTACACAGCAAGCAAAATAAGGGGCTTATGATACCTTTACCAG	511
QY	508	GGCCTCACTCCCTTCCCATGTTGTGCATGGGGCCAGAGATGAAGGAAGTGGACTAAA	567
DB	512	GGCCTCACTCCCTTCCCATGTTGTGCATGGGGCCAGAGATGAAGGAAGTGGACTAAA	571
QY	568	GTATTAACCCCTTAAGCTCCCATTTGGCTGAAGACATGAAGTACCCACCCCTTCAGAGCA	627
DB	572	GTATTAACCCCTTAAGCTCCCATTTGGCTGAAGACATGAAGTACCCACCCCTTCAGAGCA	631
QY	628	GGTCTCATTTGAAGGGAGGGCTTACTGCTTGTGTAAATTAAGATTTTACAGACCA	686
DB	632	GGTCTCATTTGAAGGGAGGGCTTACTGCTTGTGTAAATTAAGATTTTACAGACCA	690

RESULT 7
ABL40810
ID ABL40810 standard; DNA; 602 BP.

AC ABL40810

DT 03-JUL-2002 (first entry)

Human tyrosine phosphatase-like enzyme related EST sequence #4

KM Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytosolatic;
KM cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant
KM cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KM antianxietytic; hypotensive; vulnerary; gene therapy; human; gene;
KM expressed sequence tag; EST; ds.

OS Homo sapiens.

PN WO200220747-A2.

PD 14-MAR-2002.

PF 05-SEP-2001; 2001WO-EP10205.
XY

PR 11-SEP-2000; 2000US-231568P.
DE 05-DEC-2000; 2000US-231403D

XX
XX
XX

(EARTH & HEAVEN)

XX
DT XXXXXX FBI, 2003-
DB

XX
XX
New business transactions

PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases -

PS Disclosure; Fig 7; 117pp; English.

The invention relates to a purified human tyrosine phosphatase-like enzyme polypeptide. The enzyme can be expressed by standard recombinant methodology. The tyrosine phosphatase-like enzyme and encoding polynucleotides are useful for screening for modulators which are used for treating a tyrosine phosphatase-like enzyme dysfunction related disease such as a central nervous system (CNS) disorder, diabetes, obesity, chronic obstructive pulmonary disease, cardiovascular disease, cancer, anorexia, cachexia, wasting disorders, appetite suppression, and other eating disorder such as bulimia. CNS disorders include brain injuries, Parkinson's disease, dementia, multiple sclerosis, stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human immunodeficiency virus (HIV) dementia, pain associated with CNS disorders and cardiovascular diseases include myocardial infarction, ischemic diseases of the heart, atrial and ventricular arrhythmia, hypertensive vascular diseases and peripheral vascular diseases. The enzyme is useful in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases or abnormalities related to the presence of mutations in the encoding nucleic acid sequences. Sequences AB140807-813 represent human tyrosine phosphatase-like enzyme related expressed sequence tag (EST) fragments.

SQ Sequence 602 BP; 109 A; 202 C; 179 G; 110 T; 2 other,

Query Match	82.6%	Score 583.8	DB 24	Length 602
Best Local Similarity	99.3%	Pred. No. 9.5e-113		
Matches 596	Conservative 0	Mismatches 3	Indels 1	Gaps 1

QY	42	GGCCCCAACTTCCTCCCTGGGGGCTCTCCGGGGCCGGCTGGCGGAACTGGCGTGGCGGGGCT	101
Db	3	GGCGCCCACTTCCTCGGGGTCCTCCGGGGCCGGCTGGCGGAACTGGCGTGGCGGGGCT	62
QY	102	CCCCGCCCACTACCAATTCCTGTTGGACATGGGGCGTGGCGGCACTGGTGTCCCTGACGA	161
Db	63	CCCCGCCCACTACCAATTCCTGTTGGACACTGGGGGTGGCGGCACTGGTGTCCCTGACGA	122
QY	162	GGCGGGGGCCCCCTACACGCGACAGCTGCCCGGCTACCCCTGCACCGCTGGCGATCC	221
Db	123	GGCGGGGGCCCCCTACACGCGACAGCTGCCCGGCTACCCCTGCACCGCTGGCGATCC	182
QY	222	CGAATTCGTGCGCGGGCGGGCCCCGACAGATGGAACGGTTGTGTGCAGATGCTGGAGAGAGG	281
Db	183	CGAATTCGTGCGCGGGCGGGCCCCGACAGATGGAACGGTTGTGTGCAGATGCTGGAGAGAGG	242
QY	282	CAACGCACGGGAGAGGCTGTGGGAGTGCATGTGCTGTGGGCTTTGGCGGCACACTGGCAC	341
Db	243	CAACGCACGGGAGAGGCTGTGGGAGTGCATGTGCTGTGGGCTTTGGCGGCACACTGGCAC	302
QY	342	CATGCTGGCTGTTCCTGGTGA- GGAGCGGGGCTTGGCTGCAGAGATGCCATTGCTG	400

Db 303 CAGCTGCGCTGTACTCTGTTGAANGAGCGGNCCTGGCTGACGAGATGCCATTGCTG 362
 Qy 401 AATCCGACGACATGACGACCGGCTCATGACCTGATGAGCAGAGAAAGCACTCTCC 460
 |||||
 Db 363 AATTCGACGACATGACGACCGGCTCATGACCTGATGAGCAGAGAAAGCACTCTCC 422
 Qy 461 AGTTCTACGACGAGAAATTAAGGGGCTTATGATACCTTCTACAGAGCCCTCACTCC 520
 |||||
 Db 423 AGTTCTACGACGAGAAATTAAGGGGCTTATGATACCTTCTACAGAGCCCTCACTCC 482
 Qy 521 TTCCCTCATGTTGTCGATGGGGCCAGAGATGAAGGAAATGAGCTAAAGTATTAACCTC 580
 |||||
 Db 483 TTCCCTCATGTTGTCGATGGGGCCAGAGATGAAGGAAATGAGCTAAAGTATTAACCTC 542
 Qy 581 TACTCCCATGTTGTCGATGAAGCACTGAGTAGCCACCCCTGACAGGAGGCTGATGA 640
 |||||
 Db 543 TACTCCCATGTTGTCGATGAAGCACTGAGTAGCCACCCCTGACAGGAGGCTGATGA 602
 |||||
 RESULT 8
 ID ABL40812 standard; DNA: 553 BP.
 XX
 AC ABL40812:
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Human tyrosine phosphatase-like enzyme related EST sequence #6.
 XX
 KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cyostatic;
 KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
 KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
 KW antirhythmic; hypotensive; vulnerary; gene therapy; human; gene;
 KW expressed sequence tag; EST; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200220747-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 05-SEP-2001; 2001WO-EP10205.
 XX
 PR 11-SEP-2000; 2000US-231568P.
 XX
 PR 06-DEC-2000; 2000US-251403P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Kossida S;
 XX
 DR WPI; 2002-339803/37.
 XX
 PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
 PT which are useful for preventing, treating diabetes, obesity, cancer,
 PT cardiovascular and pulmonary diseases
 XX
 PS Disclosure; Fig 9; 117pp; English.
 XX
 PS The invention relates to a purified human tyrosine phosphatase-like
 CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
 CC methodology. The tyrosine phosphatase-like enzyme and encoding
 CC polynucleotides are useful for screening for modulators which are used
 CC for treating a tyrosine phosphatase-like enzyme dysfunction related
 CC disease such as a central nervous system (CNS) disorder, diabetes,
 CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
 CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
 CC other eating disorder such as bulimia. CNS disorders include brain
 CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
 CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
 CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
 CC immunodeficiency virus (HIV) dementia. Pain associated with CNS disorders
 CC and cardiovascular diseases include myocardial infarction, ischaemic
 CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive

CC vascular diseases and peripheral vascular diseases. The enzyme is useful
 CC in diagnostic assays for detecting diseases and abnormalities or
 CC susceptibility to diseases or abnormalities related to the presence of
 CC mutations in the encoding nucleic acid sequences. Sequences ABL40807-813
 CC represent human tyrosine phosphatase-like enzyme related expressed
 CC sequence tag (EST) fragments.
 XX
 SO Sequence 553 BP; 96 A; 186 C; 168 G; 103 T; 0 other;
 Query Match 71.9%; Score 508.6; DB 24; Length 553;
 Best Local Similarity 98.6%; Pred. No. 4.6e-97;
 Matches 545; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
 Qy 29 CGATGGGCGTGGACGAGCCCGCACTCTCTCGGAGCTCTCCGGCGGCTGGCGGACTGG 88
 Db 1 CGATGGGCGTGGACGAGCCCGCACTCTCTCGGAGCTCTCCGGCGGCTGGCGGACTGG 60
 Qy 89 CGCTGCGCGGCTGCCCGCCACATGACGATGCTGTTGACGACGAGGCGTGGCGGACTGG 148
 Db 61 CGCTGCGCGGCTGCCCGCCACATGACGATGCTGTTGACGAGCTGGCGGCTGGCGGACTGG 120
 Qy 149 TGTCCCTGAC--GGAGCGGCGGCGCCCTCAGAGGACAGTGGCCGGCTTACCTGCA 206
 |||||
 Db 121 TGTCCCTGACGATGAGCTGCGGCGGCGCCCTCAGAGGACAGTGGCCGGCTTACCTGCA 180
 Qy 207 CCGCTGCGGATGCC--GACTTGTGCCGCGGCGGCGCCCGACAGATGACGCGGCTTGTGC 265
 |||||
 Db 181 CCGCTGCGGATGCCGCTGATCTTGTGCGCGGCGGCGCCCGACAGATGACGCGGCTTGTGC 240
 Qy 266 AGATGCTGACGA--GGCCAAAGCAGCGGAGAGGCTGGAGTGCATGCTGCTGGGC 324
 |||||
 Db 241 AGATGCTGACGAATGGCCAAAGCAGCGGAGAGGCTGGAGTGCATGCTGCTGGGC 300
 Qy 325 TTGGGCGGCACTGGCAGCATGCTGGCTGTACTGTGTAAGAGCGGCGCTTGGCTGCA 384
 |||||
 Db 301 TTGGGCGGCACTGGCAGCATGCTGGCTGTACTGTGTAAGAGCGGCGCTTGGCTGCA 360
 Qy 385 GGAGATGCAATGCTGTAATCCGACGACTGACGAGCCGCTCCATGAGAGCTTATGACAG 444
 |||||
 Db 361 GGAGATGCAATGCTGTAATCCGACGACTGACGAGCCGCTCCATGAGAGCTTATGACAG 420
 Qy 445 GAGAAAGCACTCTCCAGTTCTACGAGCAAGAAATTAAGGGGCTTATGACCTTCTAC 504
 |||||
 Db 421 GAGAAAGCACTCTCCAGTTCTACGAGCAAGAAATTAAGGGGCTTATGACCTTCTAC 480
 Qy 505 CAGGCCCTCACTCCCTTCCCATGTTGTGATGGGGCCAGAGTGAAGGAAAGTGGACT 564
 |||||
 Db 481 CAGGCCCTCACTCCCTTCCCATGTTGTGATGGGGCCAGAGTGAAGGAAAGTGGACT 540
 Qy 565 AAAGTATTAACG 577
 |||||
 Db 541 AAAGTATTAACG 553
 |||||
 RESULT 9
 ID ABL40808 standard; DNA: 493 BP.
 XX
 AC ABL40808:
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Human tyrosine phosphatase-like enzyme related EST sequence #2.
 XX
 KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cyostatic;
 KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
 KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
 KW antirhythmic; hypotensive; vulnerary; gene therapy; human; gene;
 KW expressed sequence tag; EST; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200220747-A2.

CC blood-related disorders (e.g., anemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
S0 Sequence 559 BP; 90 A; 192 C; 169 G; 101 T; 7 other;

Query Match 66.7%; Score 471.8; DB 24; Length 559;
Best Local Similarity 99.6%; Pred. No. 2.2e-89;
Matches 473; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 GTGGCCCTTCCAGGATGGGCGTGGACGCCCCCAACTTCCTGGGCTTCCGGCGG 74
DB 72 GTCCGCTTCCAGGATGGGCGTGGACGCCCCCAACTTCCTGGGCTTCCGGCGG 131
OY 75 GCTGGCGGAGCTGGCGCTGCGCGGCTCCCGCCCACTACAGTTCCTGTTGGACCTGG 134
DB 132 GCTGGCGGAGCTGGCGGCTGCGCGGCTCCCGCCCACTACAGTTCCTGTTGGACCTGG 191
OY 135 CGTGGCGGAGCTGGCTGTCTCTGACGAGCGGGGCGCCCTTCACAGCAGCTGCCGG 194
DB 192 CGTGGCGGAGCTGGCTGTCTCTGACGAGCGGGGCGCCCTTCACAGCAGCTGCCGG 251
OY 195 CCTACCCCTGACCGCCCTGGCGATCCCGACTTCCTGGCGGGCGCCCGACCAATGA 254
DB 252 CCTACCCCTGACCGCCCTGGCGATCCCGACTTCCTGGCGGGCGCCCGACCAATGA 311
OY 255 CCGCTTCTGAGATCGTGAAGAGCCCAAGCAGCAGGGGAGAGGCTGTGGAGTGCATG 314
DB 312 CCGCTTCTGAGATCGTGAAGAGCCCAAGCAGCAGGGGAGAGGCTGTGGAGTGCATG 371
OY 315 TGGCTTGGGCTTGGCGGCTGACGACATGCTGGCTTCTTACCTGGTGAAGAGCGGG 374
DB 372 TGGCTTGGGCTTGGCGGCTGACGACATGCTGGCTTCTTACCTGGTGAAGAGCGGG 431
OY 375 CTGGCTGACGAGATGCTTGAATCGAGATGACGACCGCGGCTCCATGAGAGC 434
DB 432 CTGGCTGACGAGATGCTTGAATCGAGATGACGACCGCGGCTCCATGAGAGC 491
OY 435 CTATGAGCAGAGAAAGAGTCTTCACTTACAGGAAAGAAATGAAGGGCC 489
DB 492 CTATGAGCAGAGAAAGAGTCTTCACTTACAGGAAAGAAATGAAGGGCC 546

RESULT 11
ABL40813/c
ID ABL40813 standard; DNA; 448 BP.
XX
AC ABL40813;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human tyrosine phosphatase-like enzyme related EST sequence #7.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antiarhythmic; hypotensive; vulnary; gene therapy; human; gene;
XX expressed sequence tag; EST; ds.
XX
OS Homo sapiens.
XX
PN W0200220747-A2.

XX
PD 14-MAR-2002.
XX
PF 05-SEP-2001; 2001WO-EP10205.
XX
PR 11-SEP-2000; 2000US-231568P.
XX 06-DEC-2000; 2000US-251403P.
XX (FARB) BAYER AG.
XX
XX Kossida S;
XX
XX WPI; 2002-339803/37.
XX
PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases
XX
PS Disclosure; Fig 10; 117p; English.
XX
CC The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. Sequences ABL40807-813
CC represent human tyrosine phosphatase-like enzyme related expressed
CC sequence tag (EST) fragments.
XX
S0 Sequence 448 BP; 93 A; 132 C; 120 G; 103 T; 0 other;

Query Match 63.1%; Score 446; DB 24; Length 448;
Best Local Similarity 100.0%; Pred. NO. 5.1e-84;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 CCGACGATCGACCGCTTGTGAGATCGTGAAGAGCCCAAGCAGCGGAGAGGCT 300
DB 448 CCGACGATCGACCGCTTGTGAGATCGTGAAGAGCCCAAGCAGCGGAGAGGCT 389
OY 301 GTGGAGTGCATGTGCTGTGGCTTGGCGGCTGACGACCGACCGACCGCTGTTACCTG 360
DB 388 GTGGAGTGCATGTGCTGTGGCTTGGCGGCTGACGACCGACCGACCGCTGTTACCTG 329
OY 361 GTGAAGAGCGGGGCTTGGCTGACGAGATGCTGTAATCCGACGACTAGACACC 420
DB 328 GTGAAGAGCGGGGCTTGGCTGACGAGATGCTGTAATCCGACGACTAGACACC 269
OY 421 GGCTTCATGAGACCTATGACGAGAGAAACAGTCTTCCAGTTTACAGAGAA 480
DB 268 GGCTTCATGAGACCTATGACGAGAGAAACAGTCTTCCAGTTTACAGAGAA 209
OY 481 TAAAGGGGCTTATGACCTTTTACAGAGGCGCTTCCAGTTTACAGAGAGG 540
DB 208 TAAAGGGGCTTATGACCTTTTACAGAGGCGCTTCCAGTTTACAGAGAGG 149
OY 541 GCCAGAGATGAAGGAATGAGCTAAATATTAACCTTACGTCCTCCATGGCTGAAGA 600
DB 148 GCCAGAGATGAAGGAATGAGCTAAATATTAACCTTACGTCCTCCATGGCTGAAGA 89
OY 601 CACTGAAGTACGCCACCCCTGACAGGAGTCTGATTGAAGGAGGCTGTACTGCTT 660

|||||
DB 88 CACTGAGTACGACCCCTGCAGGAGCTCTGATTGAAGGAGCTTGTCTCTT 29
QY 661 GTTGAATTAATGAGTTTACGACCA 686
|||||
DB 28 GTTGAATTAATGAGTTTACGACCA 3
RESULT 12
ABLA0809
ID ABLA0809 standard; DNA; 453 BP.
XX
AC ABLA0809;
XX
DE 03-JUL-2002 (first entry)
XX
DE Human tyrosine phosphatase-like enzyme related EST sequence #3.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytosolic;
cardiac; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
antiarhythmic; hypotensive; vulnery; gene therapy; human; gene;
expressed sequence tag; EST; ds.
XX
OS Homo sapiens.
XX
PN WO200220747-A2.
XX
PD 14-MAR-2002.
XX
PF 05-SEP-2001; 2001WO-EP10205.
XX
PR 11-SEP-2000; 2000US-231568P.
PR 06-DEC-2000; 2000US-231403P.
XX
PA (FARB) BAYER AG.
PI Kossida S;
XX
DR WPI; 2002-339803/37.
XX
PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases
XX
PS Disclosure; Fig 6; 117pp; English.
XX
CC The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. Sequences ABLA0807-813
CC represent human tyrosine phosphatase-like enzyme related expressed
CC sequence tag (EST) fragments.
XX
SQ Sequence 453 BP; 67 A; 161 C; 153 G; 72 T; 0 other;

Query Match 60.7%; Score 429; DB 24; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.8e-80;

Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 CCAGCGATGGGGCTGACGACCCGCCCAACTTCTCTGAGGCTTCCGGGCCGCTGGCGGA 84
|||||
DB 25 CCAGCGATGGGGCTGACGACCCGCCCAACTTCTCTGAGGCTTCCGGGCCGCTGGCGGA 84
QY 85 CTGGGCGCTCCGCGGCTCCCGCCCACTACAGTTCTGTGGACCTGGGCGTGGCGAC 144
|||||
DB 85 CTGGGCGCTCCCGCGGCTCCCGCCCACTACAGTTCTGTGGACCTGGGCGTGGCGAC 144
QY 145 CTGGTGTCCCTGACGAGGAGCGGGCCCTCAGAGGACAGCGGCCCGGCTGACCCCTG 204
|||||
DB 145 CTGGTGTCCCTGACGAGGAGCGGGCCCTCAGAGGACAGCGGCCCGGCTGACCCCTG 204
QY 205 CACCGGCTGCGGATCCCGCACTTCTGCCCCCGGCCCGCCGACAGATCGACGCTTCTG 264
|||||
DB 205 CACCGGCTGCGGATCCCGCACTTCTGCCCCCGGCCCGCCGACAGATCGACGCTTCTG 264
QY 265 CAGATCTGTGACGAGGCCAACGACGAGGAGAGGCTGTGGAGTGACCTGTCTGGGC 324
|||||
DB 265 CAGATCTGTGACGAGGCCAACGACGAGGAGAGGCTGTGGAGTGACCTGTCTGGGC 324
QY 325 TTGGGCGGCACTGGGACCACTGCTGGCCGTCTACCTGCTGAAGAGAGCGGGCTTGGCTGA 384
|||||
DB 325 TTGGGCGGCACTGGGACCACTGCTGGCCGTCTACCTGCTGAAGAGAGCGGGCTTGGCTGA 384
QY 385 GGAGATGCCATGCTGAATTCGACGACATACGACCCGCTCCATCGAGACTATGACAG 444
|||||
DB 385 GGAGATGCCATGCTGAATTCGACGACATACGACCCGCTCCATCGAGACTATGACAG 444
QY 445 GAGAAAGCA 453
|||||
DB 445 GAGAAAGCA 453
RESULT 13
AAF63563
ID AAF63563 standard; CDNA; 715 BP.
XX
AC AAF63563;
XX
XX
DT 11-MAY-2001 (first entry)
XX
DE Murine phosphatase AA023073_m coding sequence.
XX
KW phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
schizophrenia; Hamman-Rich syndrome; ss.
XX
OS Mus sp.
XX
PN WO200112819-A2.
XX
PD 22-FEB-2001.
XX
PF 11-AUG-2000; 2000WO-US22158.
XX
PR 13-AUG-1999; 99US-0149005.
XX
PA (SUGEN-) SUGEN INC.
PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Llaubins M;
XX
DR WPI; 2001-211226/21.
XX
DR P-PSDB; AAB73211.
XX
PT New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
PS Claim 1; Fig 4; 138pp; English.

```

XX  The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase coding sequence.
CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins
CC modified by phosphorylation of serine, threonine or tyrosine residues.
CC The phosphatases are useful for treating a variety of diseases: for
CC example cancer e.g. breast, uterine, prostate, head, neck, lung
CC cancers, synovial sarcoma, renal cell carcinoma, non-small cell lung
CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
CC cancer, glioblastoma, colorectal cancer and thyroid cancer,
CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, Papillon-Lévy
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
CC Blomsted syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamartomas.
XX
S0 Sequence 715 BP; 156 A; 227 C; 198 G; 134 T; 0 other;

Query Match      56.1%; Score 396.4; DB 22; Length 715;
Best Local Similarity 83.3%; Pred. No. 1.2e-73;
Matches 480; Conservative 0; Mismatches 81; Indels 15; Gaps 2;

QY 20 CTTTCCAGGATGGGCGGCTGACGCCCCCACTTCTCTGGGTGCTTCCGGCGGCTGG 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 CTCCTCCACGATGGGCGGCTGCAACCCCACTTCTCTGGGTGCTTCCGGCGGCTGG 61

QY 80 CGGACATGGCGTCCGCGGCTCCCGCCACTACAGTCTGTTGGACGTGGCGTGGC 139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 CCGGACTGGCGTGGCGGCTCCCGCCGCTGCACTACCTCTCTGCTGGACAGGCTGGC 121

QY 140 GGCACATGGTCTCCCTGACGAGCGCGGCGCCCTACAGCAGCAGTCCCGGCTCA 199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 GGCACATGGTCTCCCTGACGAGCGCGGCGCCCTACAGCAGCAGTCTCCCGGCTCA 181

QY 200 CCGTCCAGCGGCTGGGCGGCTCCCGGCTCTGCGGCGGCGGCGGCGGCGGCGGCT 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 CGCTGACCGGATGGGCGGCTCCCTGACTTTTGGCGGCGGCGGCGGCGGCGGCT 241

QY 260 TCGTGCAGATGTGTGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 TTGTGTAAGATGTGTGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 301

QY 320 TGGGCTTTGGCGGCGGCTGACGAGTGTGCTTACCTGCTGTAAGAGCGGCGGCTGG 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 TAGGCTTTGGCGGCGGCTGACGAGTGTGCTTACCTGCTGTAAGAGCGGCGGCTGG 361

QY 380 CTGACGAGATGGCATGCTGTAATCCGACGACGAGCGGCGGCGGCGGCGGCGGCT 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 CCGGAGGAGATGGCATGCTGTAATCCGACGACGAGCGGCGGCGGCGGCGGCGGCT 421

QY 440 AGCAGGAGAAAGCATGCTTTCAGTCTACAGCGAAGCAAAATTAAGGGCGCTTACTACCT 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 AACAGGAGAAAGCGCGCTTTCAGTCTTACAGCGAAGCAAAATTAAGGGCGCTTACTACCT 470

QY 500 TCTACGAGGCGGCTTACTCCCTTCCCATGTTGTGATGGGCGGCAAGATGAAGGAAGT 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 TCAACAAACCGCGCTTCTCCCTTCCCAACTCTCTGCGGCGGAGGAGGAGGAGT 526

QY 560 GGACATTAAGTAAACCTCTAGCTCCCATTTGCT 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 527 GAACATAAGTACTGATCTTCAAGTCCCTTACT 562

RESULT 14
AB140811
ID AB140811 standard; DNA; 397 BP.
AC AB140811;
XX
XX 03-JUL-2002 (first entry)
DT
XX Human tyrosine phosphatase-like enzyme related EST sequence #5.
DE
XX

```

```

KM Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cyostatic;
KM cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KM cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KM antiarrhythmic; hypotensive; vulnerary; gene therapy; human; gene;
KM expressed sequence tag; EST; ds.
XX
XX Homo sapiens.
XX WO20020747-A2.
XX
XX 14-MAR-2002.
XX
XX 05-SEP-2001; 2001WO-Ep10205.
XX
XX 11-SEP-2000; 2000US-231568P.
XX
XX 06-DEC-2000; 2000US-251403P.
XX
XX (FARB ) BAYER AG.
XX
XX Kossida S;
XX
XX WPI: 2002-339803/37.
XX
XX New human tyrosine phosphatase-like enzyme polypeptide, regulators of
XX which are useful for preventing, treating diabetes, obesity, cancer,
XX cardiovascular and pulmonary diseases
XX
XX Disclosure: Fig 8; 117pp; English.
XX
XX The invention relates to a purified human tyrosine phosphatase-like
XX enzyme polypeptide. The enzyme can be expressed by standard recombinant
XX methodology. The tyrosine phosphatase-like enzyme and encoding
XX polynucleotides are useful for screening for modulators which are used
XX for treating a tyrosine phosphatase-like enzyme dysfunction related
XX disease such as a central nervous system (CNS) disorder, diabetes,
XX obesity, chronic obstructive pulmonary disease, cardiovascular disease,
XX cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
XX other eating disorder such as bulimia. CNS disorders include brain
XX injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
XX Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
XX Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
XX immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
XX and cardiovascular diseases include myocardial infarction, ischemic
XX diseases of the heart, atrial and ventricular arrhythmia, hypertensive
XX vascular diseases and peripheral vascular diseases. The enzyme is useful
XX in diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to diseases or abnormalities related to the presence of
XX mutations in the encoding nucleic acid sequences. Sequences AB140807-813
XX represent human tyrosine phosphatase-like enzyme related expressed
XX sequence tag (EST) fragments.
XX
S0 Sequence 397 BP; 52 A; 142 C; 135 G; 68 T; 0 other;

Query Match      53.7%; Score 379.8; DB 24; Length 397;
Best Local Similarity 99.5%; Pred. No. 3.3e-70;
Matches 381; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 CCAGGATGGGCGGCTGACGCCCCCACTTCTCTGGGTGCTTCCGGCGGCTGGCGGA 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14 CCAAGTATGGGCGGCTGACGCCCCCACTTCTCTGGGTGCTTCCGGCGGCTGGCGGA 73

QY 85 CTGGGCTGCGGCGGCTCCCGCCACTACAGGATCTGTTGAGCTGGGCGGCGGAC 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 74 CTGGGCTGCGGCGGCTCCCGCCACTACAGGATCTGTTGAGCTGGGCGGCGGAC 133

QY 145 CTGGTGTCTTGTACGAGAGCGGCGGCGGCTTACAGCAGCAGTGGCGGCTCACCTG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 134 CTGGTGTCTTGTACGAGAGCGGCGGCGGCTTACAGCAGCAGTGGCGGCTCACCTG 193

QY 205 CACGCGCTGGCGGATCCCGGAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTG 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194 CACGCGCTGGCGGATCCCGGAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTG 253

```

QY 265 CAGATCGTGAGAGGCGCAACGAGGAGAGAGCTGTGGAGTGCATCTGTCTGCGC 324
 |||
 Db 254 CAGATCGTGAGAGGCGCAACGAGGAGAGCTGTGGAGTGCATCTGTCTGCGC 313
 QY 335 TTGGCCGCGACTGCGACCACTGCTGCTGTACTGTGTGAAGAGCGGGCTTGCTGCA 384
 |||
 Db 314 TTGGCCGCGACTGCGACCACTGCTGCTGTACTGTGTGAAGAGCGGGCTTGCTGCA 373
 QY 385 GGAGATGCCATTGTGAAATCCG 407
 |||
 Db 374 GGAGATGCCATTGTGAAATCCG 396

RESULT 15

ID AAF27959 standard; DNA; 453 BP.

AC AAF27959;

DT 08-MAY-2001 (first entry)

DE Murine dual specificity phosphatase DSP-11 coding sequence.

KW Mouse; dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;

KW MAP-kinase signalling cascade; cell proliferation; cancer; allergy;

KW graft-versus-host disease; autoimmune disease; metabolic disease; ss.

OS Mus sp.

PN WO200105983-A1.

PD 25-JAN-2001.

PF 19-JUL-2000; 2000WO-US19710.

PR 20-JUL-1999; 99US-0144557.

PA (CEPT-) CEPTYR INC.

PI Lucite RM, Wel B;

DR WPI; 2001-147348/15.

DR P-PSDB; AAB35276.

PT Novel dual specificity phosphatase-11 dephosphorylating activated
 PT mitogen-activated protein kinase, is used to identify agents inhibiting
 PT the enzyme activity and modulate cell proliferation -

PS Example 3; Fig 5; 65pp; English.

CC The present invention provides the protein and coding sequences of the
 CC human dual-specificity phosphatase DSP-11. The protein is involved in the
 CC MAP-kinase signalling cascade. The sequences can be used in the diagnosis
 CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
 CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
 CC cell growth and abnormal cell proliferation. The present sequence is the
 CC mouse DSP-11 coding sequence.

XX Sequence 453 BP; 82 A; 144 C; 141 G; 86 T; 0 other;

Query Match 52.8%; Score 373; DB 22; Length 453;

Best Local Similarity 89.0%; Pred. No. 8.8e-69;

Matches 403; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 31 ATGGCGGTGACGCCCACTTCTCTGGGCTTCGCGGCGGCTGCGGACTGGCG 90
 |||

Db 1 ATGGCGGTGACGCCCACTTCTCTGGGCTTCGCGGCGGCTGCGGACTGGCG 60
 |||

QY 91 CTGCGCGGCTCCCGCCACTTACCACTTCTCTGTGGAAGCTGCGGCGCACTGGTG 150
 |||

Db 61 TTGCCCCGGCTGCGCGCGCACTACAGTCTCTGTGGAAGCTGCGGCGCACTGGTG 120
 |||

QY 151 TCCCTGACGAGCGCGGCGCCCTCACAAGCAGAGCTGCCCGGCTCACTGACCGC 210
 |||

Db 121 TCCCTGACGAGCGCGGCGCCCTCACAAGTGAAGCTGCCGCTCACGCTGACCGA 180
 |||
 QY 211 CTGCGCATCCCGACTTCTGCGCGCGCGCCCGCGAGACAGATGACCGCTTGTGAGATC 270
 |||
 Db 181 ATGCGCATCCCTGACTTTTGGCCCGCGCTCCCGGAGACAGATGACCAATTTGTGAAGATC 240
 |||
 QY 271 GTGAGCAGAGCGCAACGCGAGGAGAGGCTGTGGAGTGCATGTCCTCTGGCTTGGC 330
 |||
 Db 241 GTGAGCAGAGCGCAATGCCCCGGGAGAGGCTGTGGAGTGCATGTCCTCTGGCTTGGC 300
 |||
 QY 331 CGCACTGCGACCATGCTGGCTGTACTGTGTGAAGAGCGGGCTTGCTGCGAGAGAT 390
 |||
 Db 301 CGCACTGCGACCATGCTGAGCTGTGTGAAGAGCGGGCTTGCTGCGAGAGAT 360
 |||
 QY 391 GCCATTGCTGAATCCGAGACTACGACCCGGCTCCATCGAGACCTATGAGCAGAGAGA 450
 |||
 Db 361 GCCATTGCTGAGATCCGGGCGCTGCGACCAAGATCCATTGAGACGTAAGACAGAGAGA 420
 |||
 QY 451 GCAGTCTTCCAGTCTTACCAAGCGGAAGAAATGA 483
 |||
 Db 421 GCCGTCTTCCAGTCTTACCAAGCGGAAGAAATGA 453
 |||

Search completed: November 4, 2002, 08:02:41
 Job time : 262 secs


```
; Sequence 12, Application US/09/747,835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunru
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dumanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: (LIFE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 2936
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1601)
; US-09-747-835A-12
```

```
Query Match          7.0%; Score 49.2; DB 10; Length 2936;
Best Local Similarity 46.7%; Pred. No. 0.028;
Matches 156; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 65 TTCGGGGCCGGCTGGCGGAGACTGGCGCCGCGCTCCCGCCACTACAGTTCTGT 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 949 TGGCCGGCGGTGCTGGGGCGTCCCTGAGCTGGCGCCCGCCGACCAAGTACAGCGCTTCC 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 TGAACCTGGGCGTGGCGGACCTGTGTCTCCCTGAAGGAGCGGGCCCGCTCAGACGCA 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1009 CCGACTGGCTGGACCACTGGCTACAGCAGCGCAAGCATGGGCTGCTCAGCTTCTTCT 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 GCGGCCCGCGCCACACCTGACCGCTGGCGCATCCCGCACTTGTGCCCCCGGCCCGCCG 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1069 GCGGCCCGCTGACGCGCCCTCTACAGCTTGTGCTGGCGCTGCGCCGCGCCGACCGCTACG 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 ACCAGATCGACCGCTTGTGACAGATGTGGAGCAGGCGCAAGCAGGAGGAGAGAGCTGTGG 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1129 ACTGTGTCAACCTGGAGTGAACAGAGTGTGGCCCAACAAGACACCTGTGGTGGAGG 1188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 GAGTGCACGTGTGCTTGGGCTTGGCCGACCTGGACCATGTGGGCTGTACTGTGGA 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1189 AGGAGGTCTGGCGGATGAGATCTACCTTCCTGGGAGTGTGGCCCTGGCAGCGTGT 1248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 AGGAGCGGGGCTGTGGCTGAGAGAGATGCCATTGC 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1249 CCTGTGGCGGTGACCTACTGCGCTCATTTGC 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 5
US-10-036-342-34
; Sequence 34, Application US/10036342
; Patent No. US20020090681A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenli
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3030R1C5
; CURRENT APPLICATION NUMBER: US/10/036,342
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/112514
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113300
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113430
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114140
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115552
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/116843
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/125774
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125826
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/127035
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/127706
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/129122
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/130359
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131272
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/132371
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/132379
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/132383
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138166
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/144791
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/146970
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/162506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 09/311832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/380142
```


Db 603 TCGACACCGGCTGTGTGAGCCACTGTGTGAGAG 637

RESULT 7

US-09-917-800A-1386
Sequence 1386, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castile, Arthur
APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1386
LENGTH: 2455
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 L16764
US-09-917-800A-1386

Query Match

Best Local Similarity 47.6%; Score 44.6; DB 10; Length 2455;
Matches 131; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 95 CCGGCTCCCGCCCACTTCAAGTTCCTGTGTGACCTGGGCGGCGGCACTGTGCC 154
DB 635 CCGAGGCGGAGCGCCACCAAGAGCGGGGCTGTTCGGGCTTGAACGTGCTCGGATCA 694
QY 155 TGAAGGAGGCGGCGGCGCCCTCAAGGAGAGCTGCCCGGCTTCACTGCAACCGCTGCG 214
DB 695 TCAAGAGGCGGAGCGGCGGCGGCGGCTATGAGGCTGAGACCGGGAAGGGGAGG 754
QY 215 GCAATCCCACTTTCGCGCGGCGGCGGCGGCGGAGCATGATGACCGCTTGTGAGATGTGG 274
DB 755 GCAAGCTGCTCATCTTCACTTGGGGGCGGCGGCACTTCAAGTGTCACTGACGATG 814
QY 275 ACGAGGCGGAGGAGGAGGAGGAGTGTGAGTGTGCTGTGGGCTTTGGCGCA 334
DB 815 ACGAGGCGATCTTTCGAGGAGGAGGAGGAGGCGGCGGCGGCGGAGGAGGAGGAGGAG 874
QY 335 CTGGACCACTGCTGCGGCTTCACTGAGTGTGAGAGAG 369
DB 875 TCGACACCGGCTGTGTGAGGAGGAGGAGGAG 909

RESULT 8

US-09-815-242-7988

Sequence 7988, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7988
LENGTH: 1680
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1680)
US-09-815-242-7988

Query Match

Best Local Similarity 54.2%; Score 44.4; DB 10; Length 1680;
Matches 90; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 125 TGAAGCTGGGCGTGGGCACTGTGTCTCCCTGACGAGGCGGCGCCCTCAGAGGACA 184
DB 533 TGGGCGGCGCGCCGCGCGCTGGGCGGCGGAGCCCTGCTGTGCTGAGACATCGAGT 592
QY 185 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 244
DB 593 GCCAGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 652
QY 245 ACCAGATGACGCGCTTGTGAGATGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 290
DB 653 ACTGTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 698

RESULT 9

US-09-758-269-13
Sequence 13, Application US/09758269
Patent No. US20020104120A1
GENERAL INFORMATION:
APPLICANT: IUCHI, SATOSHI
APPLICANT: KOBAYASHI, MASATOMO
APPLICANT: SHINOZAKI, KAZUO
TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
FILE REFERENCE: 3914-3
CURRENT APPLICATION NUMBER: US/09/758,269
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: JP 2001-003476
PRIOR FILING DATE: 2001-01-11

Db 1217 CGGACTGCTTCCAGCGTTCGACGCGCACACCGGTTACCGCTTGGCCCAAGTCCCTGG 1158
QY 140 GGCACCTGGTGTCTCTGACGAGCGCGGCCCCCTCACAGCAGCTCCCGGCTCA 199
Db 1157 GCATATCAGCCCGCGTCAAGCGCAAGTACCAACCGCCCGGAGAGGCGCGACCGGC 1098
QY 200 CCCTGACCGCTTGGGCACTTCTGCCCCGCGGCCCCCGACCAAGATGACCGCT 259
Db 1097 TCGCCAGGCTATGGAATCGGCGCCCTTCCGCTCCCGCGGAGCAGCAGCAGCTTG 1038
QY 260 TCGTCAGATCTGTGACGAGCGCAACGCGGAGAGGAGCTGTGGAGTCACTGTGCTC 319
Db 1037 CCGTGAATGTGCGCAACCGCCCTCGATCGCCGCGGCGGCGCGGCGGAGCTTGGCC 978
QY 320 TGGGCTTGGC 330
Db 977 TTGAATCGTC 967

RESULT 15

US-09-759-143-399
; Sequence 399, Application US/09759143
; Patent No. US20020022248A1

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Kettler, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darlick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, AlJun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 399
; LENGTH: 298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(298)
; OTHER INFORMATION: n = A,T,C or G
US-09-759-143-399

Query Match

Best Local Similarity 49.8%; Score 42; DB 10; Length 298;

Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 125 TGAACCTGGGCGTGGGACCTGTGTCTGACGAGCGGCGGCCCTCACACGAGACA 184
Db 94 TGGGCGACGCGATGATGCGTGGCGTCCGAGATGAGCGCATGGGCTTGATGAGAC 153
QY 185 GCTGCGCCGCGCTCAACCTGACCGCGCTGCGCATCCCGACTTCTGCCCGCGGCGCCCG 244
Db 154 GCATGGGCTCGTGGAGGCGCATGGGCTCCGCGCATGAGCGCATGGGCGCGGCGCTCG 213
QY 245 ACCAGATGACCGCTTCTGAGATCGTGAAGAGCGCAACGACGAGGAGAGGCTGTGG 304
Db 214 ACCACATGGCTTCANCAATTGANGCATGGGCGACGACCATGAGCGCATTTGGCTTGGCG 273

QY 305 GAGTGCATGTGCTCTGTGGGCTTTGG 329
Db 274 TGGAGCNCATGGGTGCCGCGCATGGG 298

Search completed: November 4, 2002, 10:23:40
Job time : 65 secs

Db 320 GGGACACCTGCGCCGCTGCGCCGCGGCTGATCCGTCAGGCGCTCTCTGC 379.
Qy 176 ACAGCGAGCTGCGCCGCGCTGACCTGACCGCCCTGCGATCCCGCATCTCCCGC 235
Db 380 GCGGGAGATGGACACACCGGGATGACACCCCGTGAACACCGATCCGCGCCCTGTGACCC 439
Qy 236 GCGCCCGGACGATGACGCGCTTGTGACAGATGCGAGACGAGCCACGCGAGGGAG 295
Db 440 CTCTGCGACCGACGCTCCCGGACGCGCTGCGCCCTGCGCCACGCGCCGCTCAAGGGCG 499
Qy 296 AGGCTGTGGAGTGCATGCTGCTGCGGCTTTGGCCGACATGCGC 339
Db 500 ACGACGTGTGCGCTGCGCCCTGCTGGGAGACGCGGACGACG 543

RESULT 2

US-08-482-385A-5
Sequence 5, Application US/08482385A
Patent No. 5728561

GENERAL INFORMATION:
APPLICANT: DENOVA,
TITLE OF INVENTION: GENES ENCODING BRACHED CHAIN ALPHA
TITLE OF INVENTION: KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER C. RICHARDSON
STREET: 235 EAST 42ND STREET, 20TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10017-5755

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,385A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SHEYKA, ROBERT F.
REGISTRATION NUMBER: 31,304
REFERENCE/DOCKET NUMBER: PC8346C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-573-1189
TELEFAX: 212-573-1939
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 2728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-482-385A-5

Query Match 7.8%; Score 55.2; DB 1; Length 2728;
Best Local Similarity 49.6%; Pred. No. 0.0008;
Matches 141; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 56 CCTGGGTCTTCCGGGCGGCTGCGGAGCTGGCGCTCCGCGGCTCCCGCCACTAC 115
Db 662 CCTGAGAGTGGCGCGCGCTGCTGAGAGAGCGGAGCTGGCTTCCCGACACTAC 721
Qy 116 AGTTCCTTTGAGACTGGGCGTGGGCACTGGTCTCCCTGACGAGAGCGGCGCCCTC 175
Db 722 GGGACACCTTGGCGCGCGCTGCGCGGCGCTGATCCGCTCCGAGGCTCACTCTGC 781
Qy 176 ACAGCGAGCTGCGCGCGGCTGACCTGACCGCGCTGGATCCCGCATCTTGGCGGC 235
Db 782 GCGGCGACTGGACACCGGCTACGACCCCGTGAACACCGGATGCGCGCCCTGTGACCC 841

Qy 236 GCGCCCGGACGATGACCGGCTTGTGACAGATGCTGAGACGAGCCACGCGAGGGAG 295
Db 842 CTCTGCGACCGACGCTCCCGGACGCGGCTGCGCGGCTGCGCGACCGCGCTCAAGGGCG 901
Qy 296 AGGCTGTGGAGTGCATGCTGCTGCGGCTTTGGCCGACATGCGC 339
Db 902 ACGACGTGTGCGCTGCGCCCTGCTGGGAGACGCGGACGACG 945

RESULT 3

US-08-911-853-12
Sequence 12, Application US/08911853
Patent No. 6048710

GENERAL INFORMATION:
APPLICANT: Geritise, Gijbert
APPLICANT: Quax, Wilhelmus J.
TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
TITLE OF INVENTION: EXPRESSION LEVELS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genecor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,853
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/699,092
FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gjalster, Debra J.
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC361-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7620
TELEFAX: 650-845-6504
INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 642 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-911-853-12

Query Match 6.5%; Score 46.2; DB 3; Length 642;
Best Local Similarity 48.0%; Pred. No. 0.061;
Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 4 CCGCGTGTCTGTGCGCTTCCACGAGATGGCGCTGACAGCCCGCCCACTTCTCTGGGG 63
Db 67 CTCGTGTGTCTCTATACCTTCAAGCTGAGACAGACGATGACTTCTTGGCTG 126
Qy 64 CTTCCGGGCGGCTGGGAGCTGGCGCTGCCGCGCTCCCGCCACTACCATGTTCTG 123
Db 127 CTGCGACGAGAGCGCGGACCGCGCCGCGCCAGAGAGCATGCGGAGCGGACGCGTG 186
Qy 124 TTGACCTGGGCGTGGCGGCTGCTGCTCCCTGACGAGAGCGGCGCCCTTACAGCAG 183
Db 187 TCCATTCAGCGGCTTCAACATCTGTTGCGGACAGCCCGGCGGCGGCGGCGACGAG 246
Qy 184 AGCTGCGCGGCTCACCTCCGACGCGCTGCGCATCCCGACTTCTGCGCGCGCGCC 243
Db 247 GCGCGCGCGGCGGCGGAGAGATGACCTGCTGGCGAGCTTGTCAACCGGAGCGC 306
Qy 244 GACGAGATGACCGGCTTGTGAGATGCTGAGCA 278


```

; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,853
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,092
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gjalster, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17612 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-911-853-29

Query Match          6.5%; Score 46.2; DB 3; Length 17612;
Best Local Similarity 48.0%; Pred. No. 0.14;
Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 4 CCCCCTGCTCTGTCGCTTCCAGACGATGGGCGGTGACAGCCGCCACTTCTCTGGGGT 63
DB 4085 CTCGTGGTGTCTGCTATACCTTCAGCTGACGACGAGCATGACCTTCTTGGCCTG 4026
QY 64 CTTCGGGCGGCTGCGGAGCTGGCGGCTGCCGCGCTCCCGCCACTACAGTTCTTG 123
DB 4025 CTGGCAGAGGCGCGCCAGCCGCCCGCCAGAGAGCATGGCCGAGCGCGCGT 3966
QY 124 TTGACCTGGGCGGTGGGCACTGTGTCTGACGAGGAGCGGGCCCTTCACAGCGAC 183
DB 3965 TCATATCCAGCGCTGACAGATCTGTTCGGACGCGCGGCGGCGGCGGACG 3906
QY 184 AGCTGCCCCGCTTCACCTGACCGGCTGGCGCATCCCGGACTTGTGCGCGGCGCC 243
DB 3905 GCCGCCCCCGCCAGCCGCGAGCATGACCTGTGGCGCAGCTTGTCAACCGGAGCGC 3846
QY 244 GACCAATGACCGCTTGTGCGATGCTGAGCA 278
DB 3845 AAGCGCTCCAGGCGCATCATCCAGGTGCGGCGCA 3811

RESULT 7
US-09-479-409-29/c
; Sequence 29, Application US/09479409
; Patent No. 6225106
; GENERAL INFORMATION:
; APPLICANT: Gerritse, Gijbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

```

; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479,409
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,853
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gjalster, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17612 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-479-409-29

Query Match          6.5%; Score 46.2; DB 4; Length 17612;
Best Local Similarity 48.0%; Pred. No. 0.14;
Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 4 CCCCCTGCTCTGTCGCTTCCAGACGATGGGCGGTGACAGCCGCCACTTCTCTGGGGT 63
DB 4085 CTCGTGGTGTCTGCTATACCTTCAGCTGACGACGAGCATGACCTTCTTGGCCTG 4026
QY 64 CTTCGGGCGGCTGCGGAGCTGGCGGCTGCCGCGCTCCCGCCACTACAGTTCTTG 123
DB 4025 CTGGCAGAGGCGCGCCAGCCGCCCGCCAGAGAGCATGGCCGAGCGCGCGT 3966
QY 124 TTGACCTGGGCGGTGGGCACTGTGTCTGACGAGGAGCGGGCCCTTCACAGCGAC 183
DB 3965 TCATATCCAGCGCTGACAGATCTGTTCGGACGCGCGGCGGCGGCGGACG 3906
QY 184 AGCTGCCCCGCTTCACCTGACCGGCTGGCGCATCCCGGACTTGTGCGCGGCGCC 243
DB 3905 GCCGCCCCCGCCAGCCGCGAGCATGACCTGTGGCGCAGCTTGTCAACCGGAGCGC 3846
QY 244 GACCAATGACCGCTTGTGCGATGCTGAGCA 278
DB 3845 AAGCGCTCCAGGCGCATCATCCAGGTGCGGCGCA 3811
```

```

RESULT 8
US-09-479-453-29/c
; Sequence 29, Application US/09479453
; Patent No. 6313283
; GENERAL INFORMATION:
; APPLICANT: Gerritse, Gijbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/479,453
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,853
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Debra J
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC361-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7620
TELEFAX: 650-845-6504
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 17612 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-479-453-29

Query Match 6.5%; Score 46.2; DB 4; Length 17612;
Best Local Similarity 48.0%; Pred. No. 0.14;

Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 4 CCGCGTCTCTGTCGCTTCCAGAGCATGGCGTGGAGCCGCCCACTTCTCTGGGTG 63
DB 4085 CTGCTGTGTGTGTCATCAGCTCAGCTCAGAGAGAGATGACTTCTTGGCTG 4026
QY 64 CTTCGGGCGCGGTGGCGGAGTGGCGCTGCGCGCTCCCGCCACTACAGTTCTG 123
DB 4025 CTCGCGAGCGAGCGCGCGCCAGCGCGCCAGAGAGATGCGGAGCGCGACGCTG 3966
QY 124 TTGGAGCTGGGCGTGGCGGACCTGTGTCTCTACAGAGAGCGGCGCCCTCAGCGGAC 183
DB 3965 TCCATCCAGGCGCTGAGCATCTGTTCGACAGCCCGCGCGCGGAGCGGACG 3906
QY 184 AGCTGCCGCGGCGCTCAGACCTGACCGCGCTGCGATCCCGGCTTTCGCGCGGCGCC 243
DB 3905 GCGGCGCGCGCGCGCGCGCGCGAGATGACCTTCTGCTGCTTCTCAACCGGAGCGC 3846
QY 244 GACCAAGATGACCGCTTCTGTCAGATCGTGGAGCA 278
DB 3845 AAGCGCTCCAGCGGAGATCAGGTGCGCGGCGGA 3811

RESULT 9

US-09-249-585A-2/c
Sequence 2, Application US/09249585A
Patent No. 6417002
GENERAL INFORMATION:
APPLICANT: Horlick, Robert
TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
FILE REFERENCE: 0867/0D905
CURRENT APPLICATION NUMBER: US/09/249,585A
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1926
TYPE: DNA
ORGANISM: Epstein Barr Virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1926)
OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match 6.4%; Score 45.4; DB 4; Length 1926;
Best Local Similarity 48.6%; Pred. No. 0.12;
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 19 CCTTCCAGGATGGGCGTGCAGCCGCCCACTTCTGCGGTGCTTCCGGCGGCGTG 78
DB 832 CCGCTCTGGCGCCCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 773
QY 79 GCGGAGCTGGGCGTGGCGGCGTCCCGGCGCTCCGCACTACAGTTCCTGTTGACCTGGGCGTG 138
DB 772 CCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 713
QY 139 CCGGAGCTGGTGTCCGAGAGAGAGGCGGCGCCCTCCTACAGAGAGACTGCGCGGCTC 198
DB 712 CTGCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653
QY 199 ACCCTGACCGCGCTGCGCATCCCGACTTGTGCGCGCGCGCGCGCGCGAGATGAGCGC 258
DB 652 CCGCTCTGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 593
QY 259 TTGCTGAGATCGTG 273
DB 592 CTCCTGCGCGCTGCTG 578

RESULT 10

US-09-050-863-2/c
Sequence 2, Application US/09050863
Patent No. 6114111

GENERAL INFORMATION:
APPLICANT: Lao, Ying

APPLICANT: Hwang, Betty

TITLE OF INVENTION: Mammalian Protein Interaction Cloning

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hombach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,863
FILING DATE: 30-MAR-1998

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Silva, Robin M.

REGISTRATION NUMBER: 38,304

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989

TELEFAX: (415) 949-8711

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 2580 base pairs
TYPE: nucleic acid

STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 6.4%; Score 45.4; DB 3; Length 2580;
Best Local Similarity 48.6%; Pred. No. 0.13;
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;


```

COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/620,925
FILING DATE: 21-Jul-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/910,647
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218, 002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9600 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1
Query Match      6.4% Score 45.4; DB 4; Length 9600;
Best Local Similarity 48.6%; Pred. No. 0.18;

```

QY 19 CATTTCACAGAGGAGGAGCGACAGCCGCCAATCTCTTGGGTCTTCCGGGCGGCTG 78
 1261 CCCCCTCGCCCCCTCCTCCTGCTCTGCGCCCTCTCGCCCTCTCCTCCTCCTG 1202
 QY 79 GCGGGAATGCGCGCTGCGCGGCGCTCCCGCCACTACAGATTCCTGTTGGACCTGGGCGTG 138
 Db 1201 CCCCCTCTGCTCCCGCCCTCTGCCCCCTCTGCTCTGCTGCGCCCTCCTGCTCCCTC 1142
 QY 139 CGGACACTGGTGTGCTCTGAGGAGCGGGGGGCCCTCACAAGCAGATGCGCCGGCTC 198
 1141 CTGCTCTCTGCCCTCTCTGCTCTGCGCCCTCTGCGCCCTCTGCGCCCTCTGCTCTG 1082
 QY 199 ACCCTGACCGCGCTCGCAGATCCCGAATCTGCCCCGCGGCCCCGACAGATGACCGC 258
 1081 CCCCCTCTGCTCTGCGCGCCCTCTGCCCCCTCTGCGCCCTCTGCTCTGCGCCCTCTG 1022
 QY 259 TTGCTGCAGATCGTG 273
 Db 1021 CTCTGCGCCCTCTG 1007

RESULT 15
 US-07-884-811-15/c
 ; Sequence 15, Application US/07884811
 ; Patent No. 5316921
 GENERAL INFORMATION:
 APPLICANT: Godowski1, Paul J. Lokker, Nathalie A. Mark, Melanie R.
 TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: palin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/884,811
FILING DATE: 19920518
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Gindger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-884-811-15

Query Match 6.48; Score 45.4; DB 1; Length 10596;
Best Local Similarity 48.68; Pred. No. 0.18;
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 19 CCTTCCAGGAGGAGGCGGTCAGCCGCCCAACTCTCTGCGCTCCGGGCGGCTG 78
DB 2795 CCCCCTCGCCCCCTCTCTCTCTCTGCTGCCCCCTCGCCCCCTCTCTCTCTG 2736
QY 79 GCGGAGCTGGCGGCTCCGCGCTCCGCCCACTACAGATTCTGTGGACCTGGGCGTG 138
DB 2735 CCCCCTCGCT 2676
QY 139 CGGACCTGTGTCT 198
DB 2675 CTGCT 2616
QY 199 ACCCTGACAGGCGGCTGCGGATCCCGACTTCTGCCCCGCGGCCCCGACAGATGACCGC 258
DB 2615 CCCCCT 2556
QY 259 TTCTGTGAGATCTGTG 273
DB 2555 CTCCTGCCCCCTCTCTG 2541

Search completed: November 4, 2002, 09:26:53
Job time : 138 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 12:49:58 ; Search time 33 Seconds
(without alignments)
605.685 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 150

Sequence: 1 MGVPNPFMSVLPGRLAGLA.....GSITYEQEKAVQFYQRTK 150

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	100.0	150	22	Human protein phos
2	150	100.0	150	22	Human dual specif
3	150	100.0	150	22	Human ovarian anti
4	131	87.3	150	22	Human phosphatase
5	131	87.3	150	22	Human acid sequenc
6	131	87.3	150	22	Amino acid sequenc
7	131	87.3	150	22	Human tyrosinase
8	35	23.3	150	22	Murine phosphatase
9	35	23.3	150	22	Murine dual specif
10	21	14.0	22	22	Human dual specif

ALIGNMENTS

RESULT 1
ID AAB20328
XX AAB20328 standard; Protein: 150 AA.
AC AAB20328;
XX
XX 29-MAY-2001 (first entry)
XX
XX Human protein phosphatase and kinase protein-7.
DE
XX Protein phosphatase and kinase protein; PPKP-7; human;
KW gastrointestinal disorder; immune system disorder;
KW neurological disorder; cell proliferative disorder; cancer;
KW diagnosis; therapy.
XX
XX Homo sapiens.
OS
XX
XX Key
FH Location/Qualifiers
FT 1..28
FT /label= Signal_peptide
FT 29..150
FT Protein
FT /label= Mature_protein
FT 93..105
FT /note= "tyrosine specific protein phosphatase active site signature"
FT
FT Region
FT 64..107
FT /note= "Y phosphatase signature"
FT 93..103
FT /note= "tyrosine specific protein phosphatase domain signature"
FT 126..141
FT /note= "tyrosine specific protein phosphatase domain signature"
FT 74..125
FT /note= "tyrosine specific protein phosphatase domain signature"
FT
FT Modified-site
FT 41
FT /note= "O-phosphorylated"
FT 43
FT /note= "O-phosphorylated"
FT Modified-site
FT 7
FT /note= "N-glycosylated"
XX
XX WO200120004-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000WO-US25515.
XX
XX 15-SEP-1999; 99US-0154141.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hallman JL, Baughn MR, Azimzai Y; Lu DM;
XX
XX WPI: 2001-244811/25.
XX N-PSDB; AAF30482.
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis, treatment and prevention of gastrointestinal, immune system, neurological and cell proliferative disorders -
XX
XX Claim 1; Page 91; 103pp; English.
XX
XX The present sequence is that of novel human protein phosphatase and kinase protein PPKP-7, as predicted from Incyte Clone ID No. 1606974CB1 (see AAF30482). Tissues that express PPKP-7 (as a

fraction of total tissues expressing PPKP-7) include reproductive (0.372), cardiovascular (0.140) and endocrine (0.093). Diseases or conditions associated with tissues expressing PPKP-7 (as a fraction of total tissues expressing PPKP-7) include cancer (0.558), inflammation or trauma (0.233) or cell proliferation (0.209). The encoded protein shows homology to Clon Intestinalis tyrosine phosphatase. The invention provides human PPKP-1 to -11 polypeptides (see AAB30322-32) and polynucleotides (see AAF30476-86). It also provides expression vectors, host cells, antibodies, agonists and antagonists, as well as methods for diagnosing, treating or preventing disorders associated with expression of PPKP, including gastrointestinal disorders, immune system disorders, neurological disorders and cell proliferative disorders, including cancer.

Sequence 150 AA;

Query Match 100.0%; Score 150; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 1e-145;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MGVPNPNFSWLPGRLAGLALPRLPAHQFLDLGVRHLVSLTERGPPHSDSCPLTLHR 60
1 LRIIDFCPPAPDQIDRFVQIVDEANRGEAVGHCALGFRGTMTLACTYKRGRLAGD 120
61 LRIIDFCPPAPDQIDRFVQIVDEANRGEAVGHCALGFRGTMTLACTYKRGRLAGD 120

Qy 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150

Db 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150

RESULT 2
AAB35275
ID AAB35275 standard; Protein: 150 AA.
XX AAB35275;
AC AAB35275;
XX 08-MAY-2001 (first entry)
XX Human dual specificity phosphatase DSP-11.
XX Human dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
XX MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
XX graft-versus-host disease; autoimmune disease; metabolic disease.
XX Homo sapiens.
XX WO200105983-A1.
XX 25-JAN-2001.
XX 19-JUL-2000; 2000WO-US19710.
XX 20-JUL-1999; 990S-0144557.
XX (CEPT-) CEPTYR INC.
XX Luche RM, Wei B;
XX WPI; 2001-147348/15.
XX N-PSDB; AAF27958.
XX Novel dual specificity phosphatase-11 dephosphorylating activated
XX mitogen-activated protein kinase, is used to identify agents inhibiting
XX the enzyme activity and modulate cell proliferation -
XX Claim 1; Fig 2; 65pp; English.

XX The present invention provides the protein and coding sequences of the
XX human dual-specificity phosphatase DSP-11. The protein is involved in the
XX MAP-kinase signalling cascade. The sequences can be used in the diagnosis

and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host disease, autoimmune diseases, allergies, metabolic diseases, abnormal cell growth and abnormal cell proliferation. The present sequence is the human DSP-11 protein.

Sequence 150 AA;

Query Match 100.0%; Score 150; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 1e-145;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGVPNPNFSWLPGRLAGLALPRLPAHQFLDLGVRHLVSLTERGPPHSDSCPLTLHR 60
1 LRIIDFCPPAPDQIDRFVQIVDEANRGEAVGHCALGFRGTMTLACTYKRGRLAGD 120
61 LRIIDFCPPAPDQIDRFVQIVDEANRGEAVGHCALGFRGTMTLACTYKRGRLAGD 120

Db 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150

Db 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150

RESULT 3
ABP43078
ID ABP43078 standard; Protein: 179 AA.
XX ABP43078;
XX 22-AUG-2002 (first entry)
XX Human ovarian antigen HVCAJ27, SEQ ID NO:4210.
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive.
XX Homo sapiens.
XX WO200200677-A1.
XX 03-JAN-2002.
XX 07-JUN-2001; 2001WO-US18569.
XX 07-JUN-2000; 2000US-209467P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
XX N-PSDB; ABQ56135.
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g.
XX ovarian cancer), immune disorders, cardiovascular disorders and
XX neurological diseases -
XX Claim 11; SEQ ID NO 4210; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to

recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovarian and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation), polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention. CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 179 AA;

Query Match 100.0%; Score 150; DB 23; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQPPNFSWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPHSDSCGTLIHR 60
DB 30 MGQPPNFSWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPHSDSCGTLIHR 89
QY 61 LRIIDFCPPAPQIDIRFVOIVDEANARGEAVGVHCAIGFRTGTMLACYLVKRGGLAAGD 120
DB 90 LRIIDFCPPAPQIDIRFVOIVDEANARGEAVGVHCAIGFRTGTMLACYLVKRGGLAAGD 149
QY 121 AIAEIRLRPGSIETYEOKAVFOYORK 150
DB 150 AIAEIRLRPGSIETYEOKAVFOYORK 179

RESULT 4
AAB73231
ID AAB73231 standard; Protein; 150 AA.

AC AAB73231;

DT 11-MAY-2001 (first entry)

DE Human phosphatase BAA91172_h.

KX phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizoprenia; hamartoma.

OS Homo sapiens.

PN MO200112819-A2.

PD 22-FEB-2001.

PF 11-AUG-2000; 2000WO-US22158.

PR 13-AUG-1999; 9905-0149005.

PA (STGE-) SOGEN INC.

XX XX

PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX MPI: 2001-211226/21.
DR N-PSDB; AAE63563.

XX New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizoprenia, cardiac
PS dysfunction and/or vascular disorders
XX Claim 6; Fig 5; 138pp; English.

XX The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase. Phosphatases are
CC enzymes that catalyse the dephosphorylation of proteins modified by
CC phosphorylation of serine, threonine or tyrosine residues. The
CC phosphatases are useful for treating a variety of diseases: for example
CC cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,
CC synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
CC Zonana syndrome, schizoprenia and hamartomas.

SQ Sequence 150 AA;

Query Match 87.3%; Score 131; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQPPNFSWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPHSDSCGTLIHR 60
DB 1 MGQPPNFSWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPHSDSCGTLIHR 60
QY 61 LRIIDFCPPAPQIDIRFVOIVDEANARGEAVGVHCAIGFRTGTMLACYLVKRGGLAAGD 120
DB 61 LRIIDFCPPAPQIDIRFVOIVDEANARGEAVGVHCAIGFRTGTMLACYLVKRGGLAAGD 120
QY 121 AIAEIRLRPG 131
DB 121 AIAEIRLRPG 131

RESULT 5

AAG67455
ID AAG67455 standard; Protein; 150 AA.

AC AAG67455;

DT 26-NOV-2001 (first entry)

DE Amino acid sequence of a human polypeptide.

KX Human: protein kinase; protein phosphatase; signal transduction;
KW intracellular signalling pathway.

OS Homo sapiens.

PN WO200109345-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-JP05060.

PR 29-JUL-1999; 99JP-0248036.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PA (HELI-) HELIX RES INST.

XX XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
 PI Senoo C, Nezu J;
 DR WPI; 2001-564736/63.
 XX
 PT New genes encoding protein kinase and protein phosphatase, useful for
 PT identifying modulators which can be used to treat human or animal
 PT disorders associated with the expression or function of these enzymes -
 XX
 PS Example 4; Page 297-298; 336pp; Japanese.
 XX
 CC The specification describes human protein kinase/protein phosphatases.
 CC The polypeptides are expected to participate in signal transduction
 CC in cells. The kinase phosphatases are connected with intracellular
 CC signalling pathways. Antisense oligonucleotides and compounds
 CC identified by screening (agonists or antagonists) can be used to
 CC treat human or animal disorders associated with the expression
 CC or function of the protein. In addition, the polypeptides may be used
 CC as target molecules for drug development. The present sequence
 CC represents a polypeptide, used in the course of the invention.
 CC
 XX
 SQ Sequence 150 AA;
 Query Match 87.3%; Score 131; DB 22; Length 150;
 Best Local Similarity 100.0%; Pred. No. 3.2e-126; Mismatches 0; Gaps 0;
 Matches 131; Conservative 0; Indels 0; Gaps 0;
 QY 1 MGVPNPNFVWLPGRAGLALPRPAHYQFLDLGVRLVSLTERGPPHSDSCPELTILHR 60
 DB 1 MGVPNPNFVWLPGRAGLALPRPAHYQFLDLGVRLVSLTERGPPHSDSCPELTILHR 60
 QY 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 DB 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 QY 121 AIAEIRRLRPG 131
 DB 121 AIAEIRRLRPG 131
 RESULT 6
 AAG67634
 ID AAG67634 standard; Protein: 150 AA.
 XX
 AC AAG67634;
 XX
 DT 26-NOV-2001 (first entry)
 XX
 DE Amino acid sequence of a human protein.
 XX
 KW Human; protein kinase; protein phosphatase; signal transduction.
 KW
 OS Homo sapiens.
 XX
 PN WO200109316-A1.
 XX
 PD 08-FEB-2001.
 XX
 PF 28-JUL-2000; 2000WO-JP05061.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 18-OCT-1999; 99US-0159590.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 17-FEB-2000; 2000US-0183322.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
 PI Senoo C, Nezu J;

XX
 DR WPI; 2001-570286/64.
 XX
 PT New genes encoding proteins with protein kinase/protein phosphatase
 PT activity, useful in the diagnosis and treatment of diseases -
 XX
 PS Example 4; Page 197-198; 233pp; Japanese.
 XX
 CC The specification describes human protein kinase/protein phosphatases.
 CC It is expected that the protein kinase/protein phosphatase gene
 CC participates in signal transduction in cells. The protein
 CC kinase/protein phosphatase polypeptides and polynucleotides are
 CC useful for developing diagnostics and treatment agents for human
 CC and animal diseases. The protein kinase/protein phosphatase polypeptides
 CC are useful as target molecules in designing novel drugs. The protein
 CC kinase/protein phosphatase polynucleotides are useful as a source of
 CC probes and primers, which may be used to isolate homologous sequences.
 CC The present sequence represents a human protein, which is used in the
 CC course of the invention.
 CC
 XX
 SQ Sequence 150 AA;
 Query Match 87.3%; Score 131; DB 22; Length 150;
 Best Local Similarity 100.0%; Pred. No. 3.2e-126; Mismatches 0; Gaps 0;
 Matches 131; Conservative 0; Indels 0; Gaps 0;
 QY 1 MGVPNPNFVWLPGRAGLALPRPAHYQFLDLGVRLVSLTERGPPHSDSCPELTILHR 60
 DB 1 MGVPNPNFVWLPGRAGLALPRPAHYQFLDLGVRLVSLTERGPPHSDSCPELTILHR 60
 QY 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 DB 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 QY 121 AIAEIRRLRPG 131
 DB 121 AIAEIRRLRPG 131
 RESULT 7
 ABB07846
 ID ABB07846 standard; Protein: 150 AA.
 XX
 AC ABB07846;
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Human tyrosine phosphatase-like enzyme.
 XX
 KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
 KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
 KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
 KW antiarrhythmic; hypotensive; vulnerary; gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200220747-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 05-SEP-2001; 2001WO-EP10205.
 XX
 PR 11-SEP-2000; 2000US-231568P.
 PR 06-DEC-2000; 2000US-251403P.
 XX
 PA (FARB) BAYER AG.
 XX
 FT Key Location/Qualifiers
 FT 92..108 Region
 FT /note- "the encoding nucleotide fragment for the above
 FT residues is not indicated in the corresponding
 FT DNA sequence"

PI Kossida S;
XX
DR WPI; 2002-339803/37.
DR N-PSDB; ABL40806.
XX
PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases
XX
PS Claim 18; Fig 2; 117pp; English.
XX
CC The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischaemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. The present sequence
CC represents the human tyrosine phosphatase-like enzyme polypeptide.
XX
SQ Sequence 150 AA;
Query Match 87.3%; Score 131; DB 23; Length 150;
Best Local Similarity 100.0%; Pred. No. 3, 2e-126;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVOPPNFSWVLPGRLAGLALPRLPAHYQFLDLGVNHLVSLTERGPHSDSCGLTLHR 60
DB 1 MGVOPPNFSWVLPGRLAGLALPRLPAHYQFLDLGVNHLVSLTERGPHSDSCGLTLHR 60
QY 61 LRIPLDFCPAPDQIDRFVQIVDEANARGEAVGVHICALGFRTGTMACYLVKERGLAAGD 120
DB 61 LRIPLDFCPAPDQIDRFVQIVDEANARGEAVGVHICALGFRTGTMACYLVKERGLAAGD 120
QY 121 AIAEIRRLRPG 131
DB 121 AIAEIRRLRPG 131
RESULT 8
AAB73211
ID AAB73211 standard; Protein; 150 AA.
XX
AC AAB73211;
XX
DT 11-MAY-2001 (first entry)
XX
DE Murine phosphatase AA023073_m.
XX
KW phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamartoma.
XX
OS Mus sp.
XX
PN WO200112819-A2.
XX
PD 22-FEB-2001.
XX

PF 11-AUG-2000; 2000MO-US22158.
XX
PR 13-AUG-1999; 99US-0149005.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Flouman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX
DR WPI; 2001-211226/21.
DR N-PSDB; AAF63563.
XX
PT New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
PS Claim 6; Fig 5; 138pp; English.
XX
CC The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase. Phosphatases are
CC enzymes that catalyse the dephosphorylation of proteins modified by
CC phosphorylation of serine, threonine or tyrosine residues. The
CC phosphatases are useful for treating a variety of diseases: for example
CC cancer e.g. breast, uterine, prostate, head, neck, lung cancers,
CC synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
CC Zonana syndrome, schizophrenia and hamartomas.
XX
SQ Sequence 150 AA;
Query Match 23.3%; Score 35; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 9, 6e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 IYDEANARGEAVGVHICALGFRTGTMACYLVKER 114
DB 80 IYDEANARGEAVGVHICALGFRTGTMACYLVKER 114
RESULT 9
AAB35276
ID AAB35276 standard; Protein; 150 AA.
XX
AC AAB35276;
XX
DT 08-MAY-2001 (first entry)
XX
DE Murine dual specificity phosphatase DSP-11.
XX
KW Mouse; dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
KW MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
KW graft-versus-host disease; autoimmune disease; metabolic disease.
XX
OS Mus sp.
XX
PN WO200105983-A1.
XX
PD 25-JAN-2001.
XX
PF 19-JUL-2000; 2000MO-US19710.
XX
PR 20-JUL-1999; 99US-0144557.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Luche RM, Wei B;
XX
DR WPI; 2001-147348/15.
DR N-PSDB; AAF27959.
XX

PT Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase, is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
XX
PS Example 3; Fig 6; 65pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human dual-specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signaling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC mouse DSP-11 protein.
XX
SQ Sequence 150 AA;

Query Match 23.3%; Score 35; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 9.6e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 IYDEANRGEAVGVHCAIGFGRTGTMACYLYKER 114
|||
DB 80 IYDEANRGEAVGVHCAIGFGRTGTMACYLYKER 114

RESULT: 10
AAB35273
ID AAB35273 standard; peptide; 21 AA.
XX
AC AAB35273;
XX
DT 08-MAY-2001 (first entry)
XX
DE Human dual specificity phosphatase DSP-11 active site domain.
XX
KM Human: dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
KM MAP-kinase signaling cascade; cell proliferation; cancer; allergy;
KM graft-versus-host disease; autoimmune disease; metabolic disease.
XX
OS Homo sapiens.
XX
PN WO200105983-A1.
XX
PD 25-JAN-2001.
XX
PF 19-JUL-2000; 2000WO-US19710.
XX
PR 20-JUL-1999; 99US-0144557.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Luche RM, Wei B;
XX
DR WPI; 2001-147348/15.
XX
PT Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase, is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
PS Disclosure; Page 8; 65pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human dual-specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signaling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC human DSP-11 active site domain.
XX
SQ Sequence 21 AA;

Query Match 14.0%; Score 21; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 91 VGVHCAIGFGRTGTMACYLY 111
|||
DB 1 VGVHCAIGFGRTGTMACYLY 21

Search completed: November 4, 2002, 12:57:59
Job time : 33 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 2002, 07:55:42 ; Search time 2068 Seconds

(without alignments)
5536.852 Million cell updates/sec

Title: US-09-619-380-1

Perfect score: 707
Sequence: 1 tgaaccgcgtctctgtgcc.....aaaaaaaaaaaaaaaaaaaaa 707

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	694.2	98.2	808	12	BG823348
2	690.8	97.7	709	14	BM676758
3	683.8	96.7	740	13	BT761974
4	676.8	95.7	747	14	BO679163
5	675.8	95.7	747	14	BO684685
6	675.4	95.3	1127	14	BM924391

7	671.6	95.0	709	10	BE263806	BE263806	601194169
8	645.8	91.3	725	13	BI598433	BI598433	603246943
9	621.4	87.9	904	14	BO892874	BO892874	AGENCOURT
10	608.6	86.1	712	12	BF970174	BF970174	602273385
11	592.6	83.8	828	12	BG831257	BG831257	602766474
12	587.4	83.1	639	12	BG823185	BG823185	602726543
13	583.8	82.6	602	9	AI681916	AI681916	tx50f02.x
14	555.6	78.6	931	14	BQ214823	BQ214823	AGENCOURT
15	551	77.9	584	13	BI761318	BI761318	603044078
16	542.8	76.8	634	13	BI116133	BI116133	602866343
17	508.8	72.0	1059	14	BO80732	BO80732	AGENCOURT
18	508.6	71.9	553	10	BE264444	BE264444	601191757
19	500.2	70.7	1066	14	BO682193	BO682193	AGENCOURT
20	495.4	70.1	1101	13	BI765135	BI765135	603051358
21	489	69.2	604	12	BG705013	BG705013	602866517
22	487.8	69.0	1008	14	BO892811	BO892811	AGENCOURT
23	486.4	68.8	680	13	BI765377	BI765377	603050403
24	486.2	68.8	493	10	BE220956	BE220956	hu02b01.x
25	484	68.5	485	9	AI739584	AI739584	w135404.x
26	477.8	67.6	481	9	AI302179	AI302179	GR58606.x
27	476.4	67.4	500	14	BM853970	BM853970	K-EST0136
28	475.4	67.2	477	9	AI597715	AI597715	tu51908.x
29	472.2	66.8	477	9	AI689729	AI689729	tx30f02.x
30	462	65.3	1255	14	BO678493	BO678493	AGENCOURT
31	454.8	64.3	650	12	BE728244	BE728244	601563392
32	453.8	64.2	457	12	BF062579	BF062579	7b61h10.x
33	451.6	63.9	458	9	AI432132	AI432132	tg77a08.x
34	446	63.1	448	9	AI363485	AI363485	qy69f11.x
35	441.8	62.5	445	9	AI309010	AI309010	q90b007.x
36	434.8	61.5	441	10	BE671361	BE671361	7e49h11.x
37	432	61.1	440	10	AM664958	AM664958	h198d06.x
38	429	60.7	453	9	AI616223	AI616223	au45910.y
39	428.2	60.6	521	13	BI847550	BI847550	469416 MA
40	414	58.6	450	14	BM179897	BM179897	UI-E-EJ0-
41	412	58.3	412	9	AA479435	AA479435	zv14e10.r
42	411.4	58.2	475	9	AI341528	AI341528	gx92d02.x
43	408.6	57.8	417	9	AA101788	AA101788	2K96c10.s
44	406.2	57.5	754	14	BO181810	BO181810	UI-H-EJ0-
45	402.8	57.0	406	9	AI223430	AI223430	q954b02.x

ALIGNMENTS

RESULT 1
LOCUS BG823348 808 bp mRNA linear EST 22-MAY-2001
DEFINITION 602726743F1 NIH_MGC.15 Homo sapiens CDNA clone IMAGE:4866058 5',
ACCESSION BG823348
VERSION BG823348.1 GI:14170935
KEYWORDS EST.
SOURCE Human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 808)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1732 row: f column: 11
High quality sequence stop: 745.
Location/Qualifiers 1..808

FEATURES
source

```
/organism="Homo sapiens"
/dn_xref="taxon:9606"
/clone="IMAGE:4866058"
/clone_id="N1H.MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB1; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      172 a      256 c      242 g      138 t
ORIGIN
Query Match      98.2% Score 694.2; DB 12; Length 808;
Best Local Similarity 98.9%; Pred. No. 6.3e-99;
Matches 699; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 TGACCCGCTGCTCTGTGCTGCTTCCACAGATGGGCGTGCAGCCGCCCACTTCCTG 60
Dn 51 TGACCCGCTGCTCTGTGCTGCTTCCACAGATGGGCGTGCAGCCGCCCACTTCCTG 110
QY 61 GTGCTTCGGGGCGGCTGGGCGGAGTGGCGCTGCCGCGCTCCCGCCCACTACAGTTC 120
Dn 111 GTGCTTCGGGGCGGCTGGGCGGAGTGGCGCTGCCGCGCTCCCGCCCACTACAGTTC 170
QY 121 CTGTTGAGCTGGGCGTGGCGGACCTGGTGTCCCTGACGAGGCGGGCGGCTTCACAGC 180
Dn 171 CTGTTGAGCTGGGCGTGGCGGACCTGGTGTCCCTGACGAGGCGGGCGGCTTCACAGC 230
QY 181 GACAGCTGCCCCGGGCTGCACCTCCAGCCGCTGGCGATCCCGACTTCGCCCGCGGCGC 240
Dn 231 GACAGCTGCCCCGGGCTGCACCTCCAGCCGCTGGCGATCCCGACTTCGCCCGCGGCGC 290
QY 241 CCCGACAGATCGACCGGCTTGTGCAATCTGTGACAGGCGCCAAACGACGGGAGAGGCT 300
Dn 291 CCCGACAGATCGACCGGCTTGTGCAATCTGTGACAGGCGCCAAACGACGGGAGAGGCT 350
QY 301 GTGGGAGTGCACCTGTGCTGTGGGCTTGGCGGACCTGGCGACATGCTGGCTGTTACTG 360
Dn 351 GTGGGAGTGCACCTGTGCTGTGGGCTTGGCGGACCTGGCGACATGCTGGCTGTTACTG 410
QY 361 GTGAAGAGCGGGGCTTGGCTGTCAGAGAGATGCCATTGCTGAATCCGACGACTACGACC 420
Dn 411 GTGAAGAGCGGGGCTTGGCTGTCAGAGAGATGCCATTGCTGAATCCGACGACTACGACC 470
QY 421 GGTCCATCGAGACTATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Dn 471 AGCTCCATCGAGACTATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 530
QY 481 TAAAGGGGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 540
Dn 531 TAAAGGGGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 590
QY 541 GCCAGAGATGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Dn 591 GCCAGAGATGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650
QY 601 CACTGAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Dn 651 CACTGAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
QY 661 GTTGAATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707
Dn 711 GTTGAATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757
RESULT 2
BM676758/c
LOCUS      BM676758      709 bp      mRNA      linear      EST 27-FEB-2002
```

```
DEFINITION      UI-E-EJ0-ahu-1-09-0-UI.s2 UI-E-EJ0 Homo sapiens cDNA clone
ACCESSION      UI-E-EJ0-ahu-1-09-0-UI 3', mRNA sequence.
VERSION      BM676758
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 709)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      9704447
COMMENT      Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msosares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
POLY(A)-Yes.
FEATURES
source
Location/Qualifiers
1..709
/organism="Homo sapiens"
/dn_xref="taxon:9606"
/clone="UI-E-EJ0-ahu-1-09-0-UI"
/clone_id="UI-E-EJ0"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site:1: EcoR I; Site:2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dfr)18 tail. The
sequence tags for this library are: fetal eyes, AGATTAACA
; lens, CCAATAGCA; eye anterior segment, AATGCCGAT;
optic nerve, CCATTAAGT; retina, CCGG; Retina Foveal and
Macular, GTCC; RPE and Choroid, ACCTA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_LIB=UI-E-EJ0
TAG_TISSUE=Foveal and Macular Retina
TAG_SEQ=GTCC"
BASE COUNT      135 a      203 c      227 g      144 t
ORIGIN
Query Match      97.7% Score 690.8; DB 14; Length 709;
Best Local Similarity 99.6%; Pred. No. 2.3e-98;
Matches 703; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 TGACCCGCTGCTCTGTGCTGCTTCCACAGATGGGCGTGCAGCCGCCCACTTCCTG 60
Dn 705 TGACCCGCTGCTCTGTGCTGCTTCCACAGATGGGCGTGCAGCCGCCCACTTCCTG 647
QY 61 GTGCTTCGGGGCGGCTGGGCGGAGTGGCGCTGCCGCGCTCCCGCCCACTACAGTTC 120
```

FEATURES	source
Db	646 GTGCTTCCGGGCGGCGTGGCGAGTGGCGCTCCGCGCTCCCGCCACTACAGTTT
Oy	121 CTGTTGGACCTGGGGCGTGGGGACCTGGTGTCCCTTACCGAGAGCGGGCCCCCTCACAGC
Db	586 CTGTGGACCTGGGGCGTGGGGACCTGGTGTCCCTTACCGAGAGCGGGCCCCCTCACAGC
Oy	181 GACAGCTGCCGGGCTTCACCCCTGCACCCCGCTGGCGATCCCGACATTCGCCCGCGGGC
Db	526 GACAGCTGCCGGGCTTCACCCCGCTGCACCCCGCTGGCGATCCCGACATTCGCCCGCGGGC
Oy	241 CCCGACCAATGCACCGCTTCGTGCAGATCTGGACGAGGCCAACCGACGGGAGAGCT
Db	466 CCCGACCAATGCACCGCTTCGTGCAGATCTGGAGAGGCCAACCGACGGGAGAGCT
Oy	301 GTGGGAGTGCACGTGCTGTGGGCTTTGGCCGCACTGGGACCATGCTGGCCCTGTACTCG
Db	406 GTGGGAGTGCACGTGCTGTGGGCTTTGGCCGCACTGGGACCATGCTGGCCCTGTACTCG
Oy	361 GTGAAAGACGGGGCTTTGGCTCGAGAGATGCGCATTCGTAAATCCGACGATACGACCC
Db	346 GTGAAAGACGGGGCTTTGGCTCGAGAGATGCGCATTCGTAAATCCGACGATACGACCC
Oy	421 GGCTCATATGACACTATATACGACGAGAAAGACATCTTCAGTTCTACAGCAAGAA
Db	286 GGCTCATATGACACTATATACGACGAGAAAGACATCTTCAGTTCTACAGCAAGAA
Oy	481 TAAGGGGCTTGTAGTACCTCTTACGAGGCGCCACATCCCTCCCATGTTTCAGTGGG
Db	226 TAAGGGGCTTGTAGTACCTCTTACGAGGCGCCACATCCCTCCCATGTTTCAGTGGG
Oy	541 GCCAGATGAAAGGAAAGTGAATAATTTAAACCTCTAGCTCCCATTTGGCTGAA
Db	166 GCCAGATGAAAGGAAAGTGAATAATTTAAACCTCTAGCTCCCATTTGGCTGAA
Oy	601 CACTGAAGTACGCCACCCCTGCAGGCAAGTCTCGATGAAAGGAGAGCTGTACTGCTT
Db	106 CACTGAAGTACGCCACCCCTGCAGGCAAGTCTCGATGAAAGGAGAGCTGTACTGCTT
Oy	661 GTTGAATAAATGATTTTACGAACCAAAAAAAAAAAAAAAAAAAAAA 706
Db	46 GTTGAATAAATGATTTTACGAACCAAAAAAAAAAAAAAAAAAAAAA 1
RESULT 3	
LOCUS	B1761974
DEFINITION	B1761974 740 bp mRNA linear EST 25-SEP-2001
ACCESSION	603048920P1 NIH_MGC_116 Homo sapiens cdna clone IMAGE:5189045 5',
VERSION	B1761974
KEYWORDS	B1761974.1 GI:15753552
SOURCE	EST.
ORGANISM	human.
ORGANISM	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 740)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.
TITLE	NH-MGC Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strusberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov plate: L1AM1472 row: h column: 06 High quality sequence stop: 720.
FEATURES	Location/Qualifiers
source	1..740

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1IMAGE:5189045"
/clone_1lib="NH_MCC_116"
/lab_host="DH10B"
/name="Organ: pooled colon, kidney, stomach; Vector:
pCMV-Sport6; site_1: NotI; site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruder (Invitrogen). Research Genetics tracking code
023. Note: this is a NH_MCC Library."

```

Query Match	Similarity	96.7%;	Score 683.8;	DB 13;	Length 740;
Best Local	Similarity	99.6%;	Pred. No. 2.7e-97;		
Matches	696;	Conservative	0;	Mismatches	2;
				Indels	1;
				Gaps	1;
QY	1	TCACCCGCTGTCCTCTGCGCCCTTTCACAGAGATGGGCGTGCACAGCCGCCCAACTTCTCTGG	60		
Db	41	TGACCCGCGTGTCTCTGTGCGCTTTCACAGAGATGGGCGTGCACAGCCGCCCAACTTCTCTGG	100		
QY	61	GTGCTTCCGGGCGCGGCTGGCGGAGTGGCGCTGCCGCGGCTCCCGCCACATACAGTTC	120		
Db	101	GTGCTTCCGGGCGCGGCTGGCGGAGTGGCGGCTGCCGCGGCTCCCGCCACATACAGTTC	160		
QY	121	CTGTTGGACCTGGGGGTGGGCGGACCTGGTGTGCTCCGAGAGGAGGCGGGGCGGCGCTCACAGC	180		
Db	161	CTGTTGGACCTGGGGGTGGGCGGACCTGGTGTGCTCCGAGAGGAGGCGGGGCGGCGCTCACAGC	220		
QY	181	GACAGCTGCCCCCGGCGCTACCTGTGACCGCGCTGCGCATCCCGAGCTTCTGCGCGCGGCG	240		
Db	221	GACAGCTGCCCCCGGCGCTACCTGTGACCGCGCTGCGCATCCCGAGCTTCTGCGCGCGGCG	280		
QY	241	CCCCACAGATCGACCGGCTTGGTCAGATCTGACGAGGCGCCAACGCAOGGGGAGAGCT	300		
Db	281	CCCCACAGATCGACCGGCTTGGTCAGATCTGACGAGGCGCCAACGCAOGGGGAGAGGT	340		
QY	301	GTGGGAGTGCACCTGTGCTCTGGGCTTGGCGGACCTGGACCATGTGCGCTTTACCTG	360		
Db	341	GTGGGAGTGCACCTGTGCTCTGGGCTTGGCGGACCTGGACCATGTGCGCTTTACCTG	400		
QY	361	GTGAAGGAGCGGGGCTTGGTCGAGGAGATGCAATGCTGTAATTCGAGCAGATACGACCC	420		
Db	401	GTGAAGGAGCGGGGCTTGGTCGAGGAGATGCAATGCTGTAATTCGAGCAGATACGACCC	460		
QY	421	GGCTCCATCGAGACCTATAGACAGAGAAACAGTCTTCAGTTTACAGGAGGAGAA	480		
Db	461	GGCTCCATCGAGACCTATAGACAGAGAAACAGTCTTCAGTTTACAGGAGGAGAA	520		
QY	481	TAAAGGCGCTTAGTACCTCTTACAGGCGCTCACTCCCTCCCTCCCATGTGTGATGGG	540		
Db	521	TAAAGGCGCTTAGTACCTCTTACAGGCGCTCACTCCCTCCCTCCCATGTGTGATGGG	580		
QY	541	GCCGAGATGGAAGGAATGACATAATTTAAACCTTAAGTCCATTGGCTGGAAGA	600		
Db	581	GCCGAGATGGAAGGAATGACATAATTTAAACCTTAAGTCCATTGGCTGGAAGA	640		
QY	601	CACGAAATGATCCACCGCTCAGGACAGTCTGTTAAAGGAGGAGGCTTGATC-TGCTT	659		
Db	641	CACGAAATGATCCACCGCTCAGGACAGTCTGTTAAAGGAGGAGGCTTGATCTTCTT	700		
QY	660	TGTTGAATTAATGATTTTACGAACCAAAAAAAAAAAAAA	698		
Db	701	TGTTGAATTAATGATTTTACGAACCAAAAAAAAAAAAAA	739		

LOCUS	BO679163	747 bp	mRNA	linear	EST 15-JUL-2002
DEFINITION	AGENCODRT_8094393 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6213855				
ACCESSION	BO679163				
VERSION	BO679163.1 GI:21791842				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCM2380 row: f column: 08 High quality sequence stop: 680.				
FEATURES	location/Qualifiers				
Source	1..747				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:6213855"				
	/clone_lib="NIH_MGC_112"				
	/tissue_type="melanotic melanoma, cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCGACG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				
BASE COUNT	153 a	235 c	227 g	132 t	
ORIGIN					
Query Match	95.7%	Score 676.8;	DB 14;	Length 747;	
Best Local Similarity	99.7%	Pred. No. 3.3e-96;			
Matches 678;	Conservative	0;	Mismatches 2;	Indels	0; Gaps 0;
QY	28	GGAGTGGGCGTGCAGACCCCGCAACTTTCCTCGGGCTTCGGGCGGGCTGGCGGACTG	87		
DB	67	GGAGTGGGCGTGCAGACCCCGCAACTTTCCTCGGGCTTCGGGCGGGCTGGCGGACTG	126		
QY	88	GGCGTGCAGCGGCTCCCGCCACTACAGTTCCTGTTGGACTGGGGGTGCGGACCTG	147		
DB	127	GCGGTGCAGCGGCTCCCGCCACTACAGTTCCTGTTGGACTGGGGGTGCGGACCTG	186		
QY	148	GTTGCTTGACGAGCGGGGCGCCCTTCACAGCGACAGCTCCCGGCTCACCTTCGAC	207		
DB	187	GTTGCTTGACGAGCGGGGCGCCCTTCACAGCGACAGCTCCCGGCTCACCTTCGAC	246		
QY	208	CGCTCGGCGATCCCGGACTTCTGCGCGCGCGCCCGACACAGATGACCGTTGTTGAG	267		
DB	247	CGCTCGGCGATCCCGGACTTCTGCGCGCGCGCCCGACACAGATGACCGTTGTTGAG	306		
QY	268	ATCGTGCAGAGGCGCAACGACGAGGGGAGAGGCTGGGAGTGCAGTGCCTGGGCTTT	327		
DB	307	ATCGTGCAGAGGCGCAACGACGAGGGGAGAGGCTGGGAGTGCAGTGCCTGGGCTTT	366		
QY	328	GGCGCATGTGCACCATCTGCTGGCTGTACTGTGGTGAAGAGAGCGGCTTGGCTGCAGGA	387		
DB	367	GGCGCATGTGCACCATCTGCTGTTACTGTGTGAAGAGAGCGGCTTGGCTGCAGGA	426		

OY	388	GATGCGATTGGCTGAAGTAATCCGACGACTACGACC	666CTCCTCATCGAGACTATGAGACGAG	447
DB	427	GATGCCATTGCCTGAAGTAATCCGACGACTACGACC	666CTCCTCATCGAGACTATGAGACGAG	486
OY	448	AAAGCAGTCTTCACATCTTACACAGGCAAGAATAA	GAAGGGCCCTTAGTACCCCTTACACG	507
DB	487	AAAGCAGTCTTCACATCTTACACAGGCAAGAATAA	GAAGGGCCCTTAGTACCCCTTACACG	546
OY	508	GGCCCTCACTCCCCCTCCCCTCATGTGTGCATATG	GGGCGACAAGATGAAGGGAATGGACTAAA	567
DB	547	GGCCCTCACTCCCCCTCCCCTCATGTGTGCATATG	GGGCGACAAGATGAAGGGAATGGACTAAA	606
OY	568	GATATTAACCCCTCTAGTCTCCCATTTGGCTTGGA	TGTAATGATGATTTTACGACCAA	627
DB	607	GATATTAACCCCTCTAGTCTCCCATTTGGCTTGGA	TGTAATGATGATTTTACGACCAA	666
OY	628	GCTCTGATTTGAAGGGGAGGCTTGATCGCTTTGT	TGTAATGATGATTTTACGACCAA	687
DB	667	GCTCTGATTTGAAGGGGAGGCTTGATCGCTTTGT	TGTAATGATGATTTTACGACCAA	726
OY	688	AAAAAAAAAAAAAAAAAAAAA	707	
DB	727	AAAAAAAAAAAAAAAAAAAAA	746	
RESULT 5	B0684685			
LOCUS	B0684685	747 bp	mRNA	linear EST 15-JUL-2002
DEFINITION	AGENCYCOURT_8061617 NIH_MGC_112 Homo sapiens CDNA clone IMAGE:6214077			
ACCESSION	B0684685			
VERSION	B0684685.1 GI:21797364			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/			
AUTHORS	1 (bases 1 to 747)			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bs-jemail.nih.gov Tissue Procurement: DCD/D/DPF CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCM2380 row: m column: 22 High quality sequence stop: 699. Location/Qualifiers 1..747 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="6214077" /clone_id="NIH_MGC_112" /tissue_type="melanotic melanoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTR7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."			
FEATURES	source			
BASE COUNT	153 a 235 c 227 g 132 t			
ORIGIN				
Query Match	95.7%	Score 676.8;	DB 14;	Length 747;
Best Local Similarity	99.7%;	Pred. No. 3.3e-96;		


```

QY 661 GTTGATATAATGACTTTTACGACC 685
DB 733 GTTGATATAATGACTTTTACCAACC 757

RESULT 7
LOCUS BE263806 709 bp mRNA linear EST 13-JUL-2000
DEFINITION 601194159p1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3537904 5',
mRNA sequence.
ACCESSION BE263806
VERSION BE263806.1 GI:9137353
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 709)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-remail.nih.gov
Plate: LICM232 row: j column: 17
High quality sequence stop: 705.

FEATURES
Source
1..709
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3537904"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site.1: XhoI; Site.2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 137 a 233 c 206 g 133 t

Query Match 95.0%; Score 671.6; DB 10; Length 709;
Best Local Similarity 98.6%; Pred. No. 2.2e-95;
Matches 688; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 TGACCCGCTGTCCTGTCGCCCTTCCACGATGGCGGCGGCCCACTTCTCTGG 60
DB 13 TGACCCGCTGTCCTGTCGCCCTTCCACGATGGCGGCGGCCCACTTCTCTGG 72

QY 61 GTGCTTCGGGGCGGCTGCGGAGTGGCGGCGGCGGCCCACTTCTCTGG 120
DB 73 GTGCTTCGGGGCGGCTGCGGAGTGGCGGCGGCGGCCCACTTCTCTGG 132

QY 121 CTGTTGACCTGCGGCTGCGGACCTGTCCTCTGACGAGCGGCGGCCCTTCAAC 180
DB 133 CTGTTGACCTGCGGCTGCGGACCTGTCCTCTGACGAGCGGCGGCCCTTCAAC 192

QY 181 GACAGCTCCCGGCGCTACCCCTGACCGCGCTGCGCATCCCGACTTCTGCGCGGCG 240
DB 193 GACAGCTCCCGGCGCTACCCCTGACCGCGCTGCGCATCCCGACTTCTGCGCGGCG 252

QY 241 CCCGACCGATGACCGCTTGTGAGATCGTGGACGAGCGCAAGCGACGAGGGAAGGCT 300
DB 253 CCCGACCGATGACCGCTTGTGAGATCGTGGACGAGCGCAAGCGACGAGGGAAGGCT 312

QY 301 GTGGAGAGCACTGCTCTGGGCTTTGGCGCACTGGACACCTGCTGGCTTTACCTG 360
DB 313 GTGGAGAGCACTGCTCTGGGCTTTGGCGCACTGGACACCTGCTGGCTTTACCTG 372

FEATURES
Source
1..725
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5294579"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
```

Query Match 91.3%; Score 645.8; DB 13; Length 725;
 Best Local Similarity 98.7%; Pred. No. 2.2e-91;
 Matches 672; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

25 CCAGCGATGGGGTGCACGCCCCCACTTCTCTGGTGTCTCCGGCCGGTGGCGGA 84
 |||
 34 CCAGCGATGGGGTGCACGCCCCCACTTCTCTGGTGTCTCCGGCCGGTGGCGGA 93
 |||
 85 CTGGCGCTGGCGGGGCTCCCGCCACTTACAGTTCCTGTGGACCTGGCGCGGCAC 144
 |||
 94 CTGGCGCTGGCGGGGCTCCCGCCACTTACAGTTCCTGTGGACCTGGCGCGGCAC 153
 |||
 145 CTGGTGTCTCTGACGAGGAGCGGGGCCCTCACAGCAGCTGCCCCGCTCACCTG 204
 |||
 154 CTGGTGTCTCTGACGAGGAGCGGGGCCCTCACAGCAGCTGCCCCGCTCACCTG 213
 |||
 205 CACCGCTGGGATCCCGCATCTCTGCCCCGGCCCGCCAGATGACCGCTTGTG 264
 |||
 214 CACCGCTGGGATCCCGCATCTCTGCCCCGGCCCGCCAGATGACCGCTTGTG 273
 |||
 265 CAGATGCTGAGCAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 324
 |||
 274 CAGATGCTGAGCAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 333
 |||
 325 TTGGCGCGCACTGGCAGCATGCTGGCTGTCTTACCTGTGTAAGAGCGGGCTTGGCTGCA 384
 |||
 334 TTGGCGCGCACTGGCAGCATGCTGGCTGTCTTACCTGTGTAAGAGCGGGCTTGGCTGCA 393
 |||
 385 GGAGATGCCATTTGCTGAATCCAGCAGCTACGACCGCGCTTCATCGAGACTATGAGCAG 444
 |||
 394 GGAGATGCCATTTGCTGAATCCAGCAGCTACGACCGCGCTTCATCGAGACTATGAGCAG 453
 |||
 445 GAGAAAGCAGTCTTCCAGTCTACGAGCAGAAAGAAATAGGGCCCTAGTACCTCTAC 504
 |||
 454 GAGAAAGCAGTCTTCCAGTCTACGAGCAGAAAGAAATAGGGCCCTAGTACCTCTAC 513
 |||
 505 CAGGCGCTCACTCCCTCCCATGTTGTCGATGAGGGCCAGAGATGAGGAGGAGTGGACT 564
 |||
 514 CAGGCGCTCACTCCCTCCCATGTTGTCGATGAGGGCCAGAGATGAGGAGGAGTGGACT 573
 |||
 565 AAGATTTAAACCTCTAGCTCCATGGCTGAAGACACTGAAGTGAAGCCACCTGACAG 624
 |||
 574 AAGATTTAAACCTCTAGCTCCATGGCTGAAGACACTGAAGTGAAGCCACCTGACAG 633
 |||
 625 GAGAGCTGATTTGAAGGAGGAGCTGTACTGCTTGTGAATAATGATTTTACGAAC 684
 |||
 634 GAGAGCTGATTTGAAGGAGGAGCTGTACTGCTTGTGAATAATGATTTTACGAAG 691
 |||
 685 CAAAAAAAAAAAAAAAAAAAAA 705
 |||
 692 CCACGACAAAAACAAAAAAA 712
 |||

RESULT 9
 B0892874 904 bp mRNA linear EST 16-AUG-2002
 LOCUS B0892874
 DEFINITION AGENCOORT.8042566 NIH_MGC_112 Homo sapiens CDNA clone IMAGE:6090341
 5', mRNA sequence.
 ACCESSION B0892874
 VERSION B0892874.1 GI:22284888
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Creniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@pds-remail.nih.gov
 Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LCM231 row: j column: 06
 High quality sequence stop: 638.

FEATURES
 source
 1. 904
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6090341"
 /clone_lib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site: 1; XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC Library."

BASE COUNT 185 a 270 c 279 g 169 t 1 others
 ORIGIN

Query Match 87.9%; Score 621.4; DB 14; Length 904;
 Best Local Similarity 96.2%; Pred. No. 1.2e-87;
 Matches 658; Conservative 0; Mismatches 22; Indels 4; Gaps 2;

28 GCGATGGGGTGGAGCCCCCACTTCTCTGGTGTCTCCGGCCGGTGGCGGACTG 87
 |||
 67 GCGATGGGGTGGAGCCCCCACTTCTCTGGTGTCTCCGGCCGGTGGCGGACTG 126
 |||
 88 GCGCTCCGCGGGTCCCGCCCACTACAGTTCCTGTGAGCTGGGCGTGGCGGACTG 147
 |||
 127 GCGCTCCGCGGGTCCCGCCCACTACAGTTCCTGTGAGCTGGGCGTGGCGGACTG 186
 |||
 148 GTGTCCCTGACGAGGCGGGGCCCTCACAGCAGACGTCGCCGCGCTCACCTGAC 207
 |||
 187 GTGTCCCTGACGAGGCGGGGCCCTCACAGCAGACGTCGCCGCGCTCACCTGAC 246
 |||
 208 CGCTCGGATGCCGCACTTCTCCCGCGGGCCCGCCAGCAATGACGCTTGTGACAG 267
 |||
 247 CGCTCGGATGCCGCACTTCTCCCGCGGGCCCGCCAGCAATGACGCTTGTGACAG 306
 |||
 268 ATCTGAGAGAGGCAACGACGAGGAGAGGCTGTGAGAGTGCAGTGTGCTGGGCTTT 327
 |||
 307 ATCTGAGAGAGGCAACGACGAGGAGAGGCTGTGAGAGTGCAGTGTGCTGGGCTTT 366
 |||
 328 GCGCGCACTGGCACATGCTGGCTGTACCTGTGAAGAGCGGGCTTGTGTCAGGA 387
 |||
 367 GCGCGCACTGGCACATGCTGGCTGTACCTGTGAAGAGCGGGCTTGTGTCAGGA 426
 |||
 388 GATGCCATGTGAAATCCGACACTACGACCCGGCTCCATGACACTATGACGAGAG 447
 |||
 427 GATGCCATGTGAAATCCGACACTACGACCCGGCTCCATGACACTATGACGAGAG 486
 |||
 448 AAGCAGCTTCCAGTTCACAGCAGCAAGAAATAAGGGGCTTATACCTCTACAG 507
 |||
 487 AAGCAGCTTCCAGTTCACAGCAGCAAGAAATAAGGGGCTTATACCTCTACAG 546
 |||
 508 GCCCTCAGTCCCTTCCCATGTTGTGATGGGGCCAGAGATGAAGGAAGTGAGCTAA 567
 |||
 547 GCCCTCAGTCCCTTCCCATGTTGTGATGGGGCCAGAGATGAAGGAAGTGAGCTAA 606
 |||
 568 GTATTAAACCTCTAGCTCCATGGCTGAAGACACTGAAGTGAAGTGAAGTGAAG 627
 |||
 607 GTATTAAACCTCTAGCTCCATGGCTGAAGACACTGAAGTGAAGTGAAGTGAAG 666
 |||
 628 GGTCC--TGATTGAAGGAGGAGGCTGTACTG--CTTGTGTAATAATGATTTACGA 683
 |||

Db 667 AGGCCCTGATTGTAAGGGAGGCTTGACTGCTTTGTTGCAATATGACTTTTACGA 726

QY 684 CCAAAAAAAAAAAAAAAAAA 707

Db 727 AACCAAGGAAAAAAAAAGAAAGAA 750

RESULT 10

BF970174 712 bp mRNA linear EST 22-JAN-2001

LOCUS 602273385F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361564 5'

DEFINITION mRNA sequence.

ACCESSION BF970174

VERSION BF970174.1 GI:12337389

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 712)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: LLM10004 row: 1 column: 21

High quality sequence stop: 657.

Location/Qualifiers

1. 712

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4361564"

/clone_lib="NIH_MGC_84"

/tissue_type="adrenal cortex carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 135 a 231 c 216 g 130 t

ORIGIN

Query Match 86.1%; Score 608.6; DB 12; Length 712;

Best Local Similarity 94.1%; Pred. No. 1.3e-85;

Matches 643; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 25 CCACGAGTGGGCGCTGACAGCCGCCCACTTCTCTGGGCTTCGGGCGGCTGCGGGA 84

Db 23 CCACGAGTGGGCGCTGACAGCCGCCCACTTCTCTGGGCTTCGGGCGGCTGCGGGA 82

QY 85 CTGGCGCTGCGGCGGCTCCCGCCACTACAGTTCCTGTTGGACTGGGCTGGGCGAC 144

Db 83 CTGGCGCTGCGGCGGCTCCCGCCACTACAGTTCCTGTTGGACTGGGCTGGGCGAC 142

QY 145 CTGGTGTCCCTGACGAGCGGCGGCGCTCCACAGCAGCTGCCCGGCTCACCTG 204

Db 143 CTGGTGTCCCTGACGAGCGGCGGCGCTCCACAGCAGCTGCCCGGCTCACCTG 202

QY 205 CACCGCTGCGGCTGCGGCTTCGCGCGGCGGCGGCGGCGGCGGCGGCTGCGGCTG 264

Db 203 CACCGCTGCGGCTGCGGCTTCGCGCGGCGGCGGCGGCGGCGGCGGCTGCGGCTG 262

QY 265 CAGATCGTGAAGGCGCAAGCAGCGGCGGAGAGGCTGTGGAGTGCATGCTCTGGCG 324

Db 263 CAGATCGTGAAGGCGCAAGCAGCGGCGGAGAGGCTGTGGAGTGCATGCTCTGGCG 322

QY 325 TTGGCCGCACTGGCACCACCTGCGCTGTTACCTGCTGGAAGGAGCGGGGCTTGCTGCA 384

Db 323 TTGGCCGCACTGGCACCACCTGCGCTGTTACCTGCTGGAAGGAGCGGGGCTTGCTGCA 382

QY 385 GGAGATGCCATTGCTGGAATCCGAGACTACGACCCGGCTCCATGAGAACCTATGACAG 444

Db 383 GGAGATGCCATTGCTGGAATCCGAGACTACGACCCGGCTCCATGAGAACCTATGACAG 442

QY 445 GAGAAAGCAGCTTTCACCTTACACGAGCAAGAAATAGGGGCTTGTACCTCTAC 504

Db 443 GAGAAAGCAGCTTTCACCTTACACGAGCAAGAAATAGGGGCTTGTACCTCTAC 502

QY 505 CAGGCGCTACCTCCCTTCCCATGTTGTCATGGGGCGGCGAGATGAAGGAAATGACT 564

Db 503 CAGGCGCTACCTCCCTTCCCATGTTGTCATGGGGCGGCGAGATGAAGGAAATGACT 562

QY 565 AAAGTATTAACCTCTGCTGCTCCATGCTGCTGGAAGACACTGAATGACCCCTGAC 624

Db 563 AAAGTATTAACCTCTGCTGCTGCTGGAAGACACTGAATGACCCCTGAC 621

QY 625 GCAGTCTGATTGAAGGAGGAGCTTGCTGCTTGTGATAATGATTTACGAC 684

Db 622 GCAGTCTGATTGAAGGAGGAGCTTGCTGCTTGTGATAATGATTTACGAC 681

QY 685 CAAAAAAAAAAAAAAAAA 707

Db 682 CACGAACCAAGCAACCAACCA 704

RESULT 11

BG831257 828 bp mRNA linear EST 22-MAY-2001

LOCUS 602766474F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4908258 5'

DEFINITION mRNA sequence.

ACCESSION BG831257

VERSION BG831257.1 GI:14178844

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 828)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: LLM1814 row: d column: 19

High quality sequence stop: 610.

Location/Qualifiers

1. 828

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4908258"

/clone_lib="NIH_MGC_42"

/tissue_type="epitheloid carcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 158 a 315 c 204 g 151 t
ORIGIN

Note: this is a NIH_MGC Library. !
Query Match 83.8%; Score 592.6; DB 12; Length 828;
Best Local Similarity 99.0%; Pred. No. 3.8e-83;
Matches 617; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

OY 25 CCAGGATGGGGGTCGACGCCCCCAACATTCCTGGGCTTCGCGGCGCGCTGGCGGA 84
DB 13 CCAGGATGGGGGTCGACGCCCCCAACATTCCTGGGCTTCGCGGCGCGCTGGCGGA 72
OY 85 CTGGGCTGCGCGGCTCCCGCCACATACAGATTCCTGGGCTGGGCGCTGGCGAC 144
DB 73 CTGGGCTGCGCGGCTCCCGCCACATACAGATTCCTGGGCTGGGCGCTGGCGAC 132
OY 145 CTGGGCTGCGCGGCTCCCGCCACATACAGATTCCTGGGCTGGGCGCTGGCGAC 204
DB 133 CTGGGCTGCGCGGCTCCCGCCACATACAGATTCCTGGGCTGGGCGCTGGCGAC 192
OY 205 CACC-GCCTGGGCTCCCGCCACATTCCTGGGCTTCGCGGCGCGCTGGCGCT 263
DB 193 CACCTGGCTGGGCTCCCGCCACATTCCTGGGCTTCGCGGCGCGCTGGCGCT 252
OY 264 GCAGATCGTGGACGAGGCGCAACGACGAGGAGGCTGGGAGTGTCTGTCTGGG 323
DB 253 GCAGATCGTGGACGAGGCGCAACGACGAGGAGGCTGGGAGTGTCTGTCTGGG 312
OY 324 CTGGGCTGCGCGGCTCCCGCCACATTCCTGGGCTGGGCGCTGGCGAC 383
DB 313 CTGGGCTGCGCGGCTCCCGCCACATTCCTGGGCTGGGAGGAGGCTGGGCT 372
OY 384 AGGAGATGCGATTCCTGGGCTCCCGCCACATTCCTGGGCTTCGCGGCTGGG 443
DB 373 AGGAGATGCGATTCCTGGGCTCCCGCCACATTCCTGGGCTTCGCGGCTGGG 432
OY 444 GGAGAAAGCAGATTCCTGGGCTCCCGCCACATTCCTGGGCTTCGCGGCTGGG 503
DB 433 GGAGAAAGCAGATTCCTGGGCTCCCGCCACATTCCTGGGCTTCGCGGCTGGG 492
OY 504 CCAGGCTGCGCGGCTCCCGCCACATTCCTGGGCTTCGCGGCTGGG 563
DB 493 CCAGGCTGCGCGGCTCCCGCCACATTCCTGGGCTTCGCGGCTGGG 552
OY 564 TAAAGTATTAACCTTCAGTCCCATTCGGAAGACAGTAAGTACCCACCCCTGCA 623
DB 553 TAAAGTATTAACCTTCAGTCCCATTCGGAAGACAGTAAGTACCCACCCCTGCA 612
OY 624 GGC-AGGTCTGATGGAAGGGA 645
DB 613 GGCAAGGTCTGATGGAAGGGA 635

RESULT 12
LOCUS BG823185 639 bp mRNA linear EST 22-MAY-2001
DEFINITION 60273654.F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4866035 5',
ACCESSION BG823185
VERSION BG823185.1 GI:14170772
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 639)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcrabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LINC1732 row: e column: 12
High quality sequence stop: 638.

FEATURES
source location/Qualifiers
1. 639

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4866035"
/clone_11b="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 106 a 220 c 200 g 113 t
ORIGIN

```

Query Match 83.1%; Score 567.4; DB 12; Length 639;
Best Local Similarity 99.8%; Pred. No. 2.8e-82;
Matches 588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGACCCGGTGTCTGGCTTCCTCCAGGATGGGCGTGGACCCCAACTTCCTGG 60
DB 51 TGACCCGGTGTCTGGCTTCCTCCAGGATGGGCGTGGACCCCAACTTCCTGG 110
OY 61 GTGCTTCGGGGCGGCTGGGCGGAGTGGCGGCGGCTCCCGCCACATTCAGTT 120
DB 111 GTGCTTCGGGGCGGCTGGGCGGAGTGGCGGCGGCTCCCGCCACATTCAGTT 170
OY 121 CTGTTGACCTGGGCGGCTGGGCGGAGTGGCGGCGGCTCCCGCCACATTCAG 180
DB 171 CTGTTGACCTGGGCGGCTGGGCGGAGTGGCGGCGGCTCCCGCCACATTCAG 230
OY 181 GACAGCTGCGGCGGCTGACCTGACCGGCGGCGGCTCCCGCCACATTCAGTT 240
DB 231 GACAGCTGCGGCGGCTGACCTGACCGGCGGCGGCTCCCGCCACATTCAGTT 290
OY 241 CCGGACAGATGACCGGCTGCGGAGATGCGGAGGCGGCGGCGGAGAGGCT 300
DB 291 CCGGACAGATGACCGGCTGCGGAGATGCGGAGGCGGCGGCGGAGAGGCT 350
OY 301 GTGGAGTGCACCTGTGCTGGGCTTTGGCGGCGGCGGCGGCGGCTTTACCTG 360
DB 351 GTGGAGTGCACCTGTGCTGGGCTTTGGCGGCGGCGGCGGCGGCTTTACCTG 410
OY 361 GTGAAGAGCGGGGCTTGGCTGCAAGATGCAATTCGTAATCCGAGCACTAGACC 420
DB 411 GTGAAGAGCGGGGCTTGGCTGCAAGATGCAATTCGTAATCCGAGCACTAGACC 470
OY 421 GGCTTCATGAGACCTATGACGAGAGAGAGAGTTCATTCATTCAGAGAGAGAA 480
DB 471 AGCTTCATGAGACCTATGAGAGAGAGAGAGTTCATTCATTCAGAGAGAGAA 530
OY 481 TAAAGGGGCTTATGATTCCTTACAGGCGCTCACTCCCTTCCCAAGTGTGATGG 540
DB 531 TAAAGGGGCTTATGATTCCTTACAGGCGCTCACTCCCTTCCCAAGTGTGATGG 590
OY 541 GCCAGAGTGAAGGAGTGAAGTAAATTAACCTCTAGAGTCCCA 589
DB 591 GCCAGAGTGAAGGAGTGAAGTAAATTAACCTCTAGAGTCCCA 639

RESULT 13
LOCUS A1681916
```

LOCUS	Al681916	602 bp	mRNA	linear	EST 16-DEC-1999
DEFINITION	tx50f02.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2273019 3'				
ACCESSION	Al681916				
VERSION	Al681916.1	GI:4892098			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 602)				
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
COMMENT	Unpublished (1997)				
CONTACT	Contact: Robert Strausberg, Ph.D.				
EMAIL	Email: cga@b-remail.nih.gov				
TISSUE	Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.				
CDNA	CDNA library Preparation: M. Bento Soares, Ph.D.				
ARRAYED	CDNA library Arrayed by: Greg Lennon, Ph.D.				
SEQUENCING	DNA Sequencing by: Washington University Genome Sequencing Center				
CLONE	Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: www.bio.lnli.gov/db/rp/image/image.html				
INSERT	Insert length: 681 std Error: 0.00				
SEG	Seq primer: -400p from Glibco				
FEATURES	High quality sequence stop: 495.				
SOURCE	Location/Qualifiers				
	1..602				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone_image="2273019"				
	/clone_lib="NCI-CGAP_Lu24"				
	/tissue_type="carcinoid"				
	/lab_host="DH10B"				
	/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldi."				
BASE COUNT	109 a 202 c 179 g 110 t 2 others				
ORIGIN					
Query Match	82.6%; Score 583.8; DB 9; length 602;				
Best Local Similarity	99.3%; Pred. No. 1e-01; 3; Indels 1; Gaps 1;				
Matches	596; Conservative 0; Mismatches 3; Indels 1; Gaps 1;				
Oy	42 GCCCCCACTTCTCTGGTGTCTCCGAGCGCGGTGGCGGGACGTGCGCGGCT 101				
Db	3 GGGCGCAACTTCTCTGGTGTCTCCGAGCGCGGTGGCGGGACGTGCGCGGCT 62				
Oy	102 CCCGCCACATACAGTTCTCTGGTGTCTGGGCGGTGGCGGACGTGTCTCTGACGGA 161				
Db	63 CCCGCCACATACAGTTCTCTGGTGTCTGGGCGGTGGCGGACGTGTCTCTGACGGA 122				
Oy	162 GCGGGGGGGGGGCTGACAGGAGAGGTGGCGGGCGGCGGCGGCGGCGGCGGATGCC 221				
Db	123 GGGGGGGGGGGGCTGACAGGAGAGGTGGCGGGCGGCGGCGGCGGCGGCGGATGCC 182				
Oy	222 GCACTTCTGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 281				
Db	183 GCACTTCTGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242				
Oy	282 CAAGCAGCGGGAGAGGCTGTGGAGTGCATGTGTCTGTGGGCTTGGCGGCGGCGG 341				
Db	243 CAAGCAGCGGGAGAGGCTGTGGAGTGCATGTGTCTGTGGGCTTGGCGGCGGCGG 302				
Oy	342 CATGTGGCGGTTCATCTGTGTGAA-GGAGCGGGGCTTGGCTGACAGAGATGCCATTGCTG 400				

Db	303	CATGCTGGCCTGTTCACGTGTCGAANGAGGCGGNGCTTGCGTCGACAGAGATGTCATTGCTG	362
QY	401	AAATCCGACGACTACGACCCGGCTCCATTCGAGACTATAGACGAGAAAGCAGTCTTC	460
Db	363	AAATCCGACGACTACGACCCGGCTCCATTCGAGACTATAGACGAGAAAGCAGTCTTC	422
QY	461	AGTTCCTACAGGAAACGAATTAAGGGGCGCTTAGTACCTTCCTACAGAGCCCTCCTCC	520
Db	423	AGTTCCTACAGGAAACGAATTAAGGGGCGCTTAGTACCTTCCTACAGAGCCCTCCTCC	482
QY	521	TTCCCATGTTGTGCATGGGGCCAGAGATGAAGGAGAGTGAAGTAAAGTATTAACCTC	580
Db	483	TTCCCATGTTGTGCATGGGGCCAGAGATGAAGGAGAGTGAAGTAAAGTATTAACCTC	542
QY	581	TAGCTCCATTGGCTGGAAGACACTGAATAGCCACCCCTGCAGGCGAGTCTGATTGA	640
Db	543	TAGCTCCATTGGCTGGAAGACACTGAATAGCCACCCCTGCAGGCGAGTCTGATTGA	602
RESULT 14	BO214823	931 bp	EST 02-MAY-2002
DEFINITION	AGNCOURT_7589286 NIH_MGC_72 Homo sapiens	cdna clone IMAGE:6063586	
ACCESSION	BO214823		
VERSION	BO214823.1	GI:20396223	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
LOCUS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 931)		
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabp-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTP CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA sequencing by: Agencourt Bioscience Corporation Gene distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LLM1336 row: 0 column: 11 High quality sequence stop: 563.		
FEATURES	location/Qualifiers		
SOURCE	1..931		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6063586"		
	/clone_1id="NIH_MGC_72"		
	/tissue_type="melanotic melanoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."		
BASE COUNT	198 a 284 c 270 g 179 t		
ORIGIN			
Query Match	78.6%; Score 555.6; DB 14; Length 931;		
Best Local Similarity	99.3%; Pred. No. 2e-77;		
Matches 558; Conservative	0; Mismatches 4; Indels 0; Gaps 0;		
QY	131	TGGCGCTGCGGACACTGTGTTCCTGACGAGCGGGGCCCCCTCAGCAGCAGCTGCC	190
Db	1	TGGCGCTGCGGACACTGTGTTCCTGACGAGCGGGGCCCCCTCAGCAGCAGCTGCC	60
QY	191	CGGGCTCAGCCTGACCGCTGCGCATCCCGAGACTTGTGCGCGGGCCCCGACACGA	250
Db	61	CGGGCTCAGCCTGACCGCTGCGCATCCCGAGACTTGTGCGCGGGCCCCGACACGA	120

251 TCAGCCGCTTCGTGACATGCTGAGACGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 310
 121 TCAGCCGCTTCGTGACATGCTGAGACGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 180
 311 ACCTGCTCTGGGCTTTGGCCGACACTGGACACTGCTGGCTGTACCTGGTGAAGAGC 370
 181 ACCTGCTCTGGGCTTTGGCCGACACTGGACACTGCTGGCTGTACCTGGTGAAGAGC 240
 371 GGGCTTGGCTGAGAGATGCTGCAATTCGAAATCGAGACAGACCGGCTCCATCG 430
 241 GGGCTTGGCTGAGAGATGCTGCAATTCGAAATCGAGACAGACCGGCTCCATCG 300
 431 AGACCTATGAGCAGAGAAAGCAGTCTCCAGTCTACAGCAGCAATTAAGGGCT 490
 301 AGACCTATGAGCAGAGAAAGCAGTCTCCAGTCTACAGCAGCAATTAAGGGCT 360
 491 TAGTACCTTCTACAGAGCCCTCACTCCCTTCCCATGTTGTGATGGGGCAGAGATG 550
 361 TAGTACCTTCTACAGAGCCCTCACTCCCTTCCCATGTTGTGATGGGGCAGAGATG 420
 551 AAGGAGAGTGAAGTAAAGTAAACCTTACCTCCATGGCTGAAGCACTGAAGTA 610
 421 AAGGAGAGTGAAGTAAAGTAAACCTTACCTCCATGGCTGAAGCACTGAAGTA 480
 611 GCCACCCCTGACGAGGCTCTGATTGAAGGGAGGCTTGTACTGTTGTGAATAA 670
 481 GCCACCCCTGACGAGGCTCTGATTGAAGGGAGGCTTGTACTGTTGTGAATAA 540
 671 TGAGTTTACGAACCAAAAAA 692
 541 TGAGTTTACGAACCAAGATA 562

RESULT 15
 LOCUS B1761318 584 bp mRNA linear EST 25-SEP-2001
 DEFINITION B1761318 mRNA sequence.
 ACCESSION B1761318
 VERSION B1761318.1 GI:15752896
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NIH-MGC http://mgc.ncl.nih.gov/
 1 (bases 1 to 584)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNI)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LILNI at:
 http://image.llnl.gov
 Plate: LILNI1460 row: k column: 20
 High quality sequence stop: 583.
 Location/Qualifiers
 1..584
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5184523"
 /clone_11b="NIH-MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site: 1; NotI; Site: 2; EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library 18

oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb.
 Insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH-MGC Library."

BASE COUNT 90 a 209 c 181 g 104 t
 ORIGIN

Query Match 77.9%; Score 551; DB 13; Length 584;
 Best Local Similarity 100.0%; Pred. No. 1.3e-76;
 Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGACCCGCTTCGTGACATGCTGAGACGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 60
 33 TCAGCCGCTTCGTGACATGCTGAGACGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 92
 61 GTGCTTCCGGGCGGGCTGGGAGACTGGGAGCTGCGGCTCCGCGCCACATCAAGTTC 120
 93 GTGCTTCCGGGCGGGCTGGGAGACTGGGAGCTGCGGCTCCGCGCCACATCAAGTTC 152
 121 CTGTTGAGCTGGGCTGGGAGACTGGGAGCTGCGGCTCCGCGCCACATCAAGTTC 180
 153 CTGTTGAGCTGGGCTGGGAGACTGGGAGCTGCGGCTCCGCGCCACATCAAGTTC 212
 181 GACAGCTGCCCGGCGGCTGACAGCTGCGGCTCCGCGCCACATCAAGTTC 240
 213 GACAGCTGCCCGGCGGCTGACAGCTGCGGCTCCGCGCCACATCAAGTTC 272
 241 CCCGACAGATCGACCGCTTGTGAGATGTTGAGAGAGGCAACGACGCGGGAGAGCT 300
 273 CCCGACAGATCGACCGCTTGTGAGATGTTGAGAGAGGCAACGACGCGGGAGAGCT 332
 301 GTGAGAGTGCATGCTGCTGAGGCTTTGGCCGACCTGGCACTGCTGGCTTTACTTG 360
 333 GTGAGAGTGCATGCTGCTGAGGCTTTGGCCGACCTGGCACTGCTGGCTTTACTTG 392
 361 GTGAGAGTGCATGCTGCTGAGGCTTTGGCCGACCTGGCACTGCTGGCTTTACTTG 420
 393 GTGAGAGTGCATGCTGCTGAGGCTTTGGCCGACCTGGCACTGCTGGCTTTACTTG 452
 421 GGCTCATCGAGACCTATGAGCAGAGAGAAAGCAGTCTTCAGTTTACAGCAGCAAGAA 480
 453 GGCTCATCGAGACCTATGAGCAGAGAGAAAGCAGTCTTCAGTTTACAGCAGCAAGAA 512
 481 TAGGGGCTTATGATCTTCTACAGGCGCTTACCTCCATGTTGTGATGGG 540
 513 TAGGGGCTTATGATCTTCTACAGGCGCTTACCTCCATGTTGTGATGGG 572
 541 GCCAGAGATGA 551
 573 GCCAGAGATGA 583

Search completed: November 4, 2002, 09:24:20
 Job time : 2075 secs

This Page Blank (uspto)